VOL. XVII APRIL, 1933 No. 2

The Journal of APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

A. M. JORDAN. Parental Occupations and Children's Intelligence Scores 103

S. L. and L. C. PRESSEY. Results of Certain Honesty Tests Given to a Group of Rural White Children and to Two Groups of Indian Children 120

FREDERICK J. GAUDET and MARCUS A. CURRY. The Effects of Business Conditions upon the Sanity of a Population 130

RALEIGH M. DRAKE. Four New Tests of Musical Talent 136

RALPH V. MERRY and FRIEDA KIEFER MERRY. The Tactual Recognition of Embossed Pictures by Blind Children 148

MARJORIE LOU PAGE. The Mechanical Ability of Subnormal Boys 164

CHARLES BIRD and F. S. BEERS. Maximum and Minimum Inner Speech in Reading 182

GRACE ARTHUR. The Predictive Value of the Kuhlmann-Binet Scale for Inmates of a State School for the Feebleminded 188

STANLEY DULSKY. Factors Influencing the Amount of Copy Read in Magazine Advertisements 195

GEORGE W. HARTMANN. The Interview as a Research and Teaching Device 205

News and Notes 212

Book Reviews 214

New Books and Pamphlets 219

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio
Printed by
C. E. PAULEY & COMPANY, INC.
225 N. New Jersey St., Indianapolis, Indiana

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry as second-class matter at the post office, Indianapolis, Indiana.

Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1933, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston, H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Wooley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

PARENTAL OCCUPATIONS AND CHILDREN'S INTELLIGENCE SCORES

A. M. JORDAN
University of North Carolina

The effect of differing environments upon mental development is an age old problem of investigation. Can man by changing the conditions under which his offspring is reared produce in them significant differences? Or does an individual of low potentialities merely neglect these surroundings rich in civilizing possibilities and react to those of the order which are suitable to the level of his inheritance? More specifically, in connection with our present study, do environments which differ widely among themselves produce in children mental changes differing in any substantial amount? The question then of whether the occupation of the parents, with its accompanying levels of leisure and refinement, influences the mental growth of children is one of general interest.

The Present Study

About 1200 school children were tested twice each year over a period of three years. The children all lived in one community and were located in Grades I to VII, inclusive. The intelligence tests used were: for Grades I and II, Pintner-Cunningham Primary Mental Tests and Dearborn Group Tests of Intelligence, Series I; for Grades III through VII, The National Intelligence Tests, Forms A and B.

Mental Status

The records of the first administration of these tests have been used to study the mental level of the pupils included in this study. The intelligence quotients scores will first be reported.

TABLE I

Classification of Intelligence Quotients of the Occupational Groups into Terman's Categories

IQ's	Mill Workers	Machinist	Merchants, Bankers, Lumbermen	Carpenter	Housekeeper	Insurance Agent	Traveling Salesmen	Clerk	Contractor	Electrician	Barber	Butcher	Professional	Painter	Stone Cutter	Trainman	Farmer	Blacksmith	Plumber	Dairyman	Laborer	Total	Non-Mill (Total)
Genius (above 139)	3							1					1									5	2
Very Superior (120-139)	13	6	6			6	8			1	2		6						1			49	36
Superior (110-119)	46	3	13	7		3	8	2	4	3	1	2	8	1		4	12			1		120	74
Normal (90-109)	241	24	22	39	5	14	22	8	12	9	12	3	19	9	7	4	37	3	4	9	7	513	272
Dull (80-89)	135	7	11	23	5	4	8	4	4	3	2	1	3	8	3		22	4	1	4	4	253	118
Borderline (70-79)	85	1	8	16		2	3	3		2	4	1	2	4	1	1	17	1		1		153	68
Feebleminded (below 70)	55	1	3	5	2		2	1	3	1				4	1		16			1	3	98	43
Total	578	42	63	90	12	29	51	19	23	19	21	7	39	26	12	9	104	8	6	16	14	1191	613
Median IQ	91	98	98	90	87	103	106	92	97	96	96	97	105	86	92	104	88	88	102	97	90		95

Table I gives the IQ distributions of the majority of pupils discussed in this paper. A few are not included because they formed groups which are entirely too small for statistical treatment. These scores are computed from the first tests which were given. It will be noted that the median scores at the foot of the table differ substantially in size. We find on the one hand that the median IQ of the children of mill workers is 91 and that of the farmers' children is 88; and, on the other, the median IQ of the children of the professional group is 105. The highest median seems to be that of the children of travelling salesmen which is 106. The median IQ's of the children of carpenters, housekeepers, painters, farmers, blacksmiths, and day laborers are 90 and below; while the median IQ's of children of insurance agents, salesmen, professions, train-men, and plumbers are 100 and above. The IQ's of the others are in between.

In Table II appear the percentages of these groups whose representatives are sufficiently large for such treatment. For the purposes of comparison the percentages which represent theoretical expectancy are placed in the last column. It may be assumed that the members of any occupational group whose percentages approximate the percentages contained in this last column are distributed normally. Any small deviation at any one level would not be a fact of much significance but any substantial deviation therefrom would at least raise questions which would need answering.

In the first place (Table II), we find the children of professional workers far ahead of any other group in the *Genius level* for in this group appear 2.5 per cent of the total number while in most of the other groups there are no representatives at all. Three out of the 578 children of mill workers appear in this group or .5 per cent. In the *very superior group* we find the children of mill workers, carpenters, and farmers below expectancy while the children of all the rest are

above. The children of the professional men, salesmen, and machinists stand out clearly both above the other groups being studied as well as above theoretical expectancy. In the *superior group* only the IQ's of the children of salesmen, merchants, lumbermen, bankers and professional men are above theoretical expectancy. The rest are substantially below. The machinists', the carpenters', and the mill workers' children are not much more than 50 per cent of the expectancy. The farmers furnish a much larger per cent of the expected than they have at any previous level. Not a single group attains the theoretical expectancy in the *normal group.* The range is from 35.0 per cent to 57.1 per cent when 60 per cent is expected. We have noted more than the theoretical expectancy in several of the categories above normal but if we look carefully we shall find a much greater group among the dull, the borderline, and the feebleminded than might have been expected.

Let us consider in some detail these groups which fall below the normal group. In the dull group 23 per cent of the children of mill workers, 25 per cent of the children of carpenters, and 21 per cent of the children of farmers fall when fifteen per cent is the theoretical expectancy. Only in the professional group with 7.7 per cent are the results much different.

When we turn to the *borderline group* the lines are even more sharply drawn. The children of the mill workers, the carpenters, and the farmers fall into this group in large numbers. Around fifteen per cent among these children according to our tests fall in the borderline group. This percentage is three times as many as theoretical expectancy.

The *feebleminded,*—those who are unable to plan and who have no ability to criticise their own acts—are much more numerous than might be expected. In this group appears a very large number of children of mill workers and of farmers. Indeed far more than chance would

TABLE II

Percentage of Each Occupational Group Appearing at Intelligence Level Indicated

IQ's	Mill Workers	Machinists	Merchants, Bankers, Lumbermen	Carpenters	Salesmen	Professional	Farmers	Total (1191)	All except Those in Mill	Theoretical Expectancy
Genius (above 139)	.5					2.5		.42	.3	.1
Very Superior (120-139)	2.2	14.3	9.5		15.7	15.4		4.16	5.8	2.4
Superior (110-119)	7.96	7.1	20.6	7.8	15.7	20.5	11.53	10.18	12.0	15.0
Normal (90-109)	41.7	57.1	35.0	43.3	43.3	48.6	35.5	43.5	44.3	60.0
Dull (80-89)	23.3	16.6	17.5	25.5	15.7	7.7	21.1	21.5	19.2	15.0
Borderline (70-79)	14.7	2.4	12.7	17.8	5.9	5.1	16.3	13.0	11.0	5.0
Feebleminded (below 70)	9.51	2.4	4.7	5.5	3.9		15.4	8.3	7.0	2.5
No. Cases	578	42	63	90	51	39	104	1191	613	

allow. These two groups stand out above the others in the percentages furnished to the feebleminded group. If the children of mill workers alone are considered, since we have here such a large sampling, nearly four times as many fall in this group as appear in the column of theoretical expectancy. In brief, these two occupational groups, the mill workers and the farmers, furnish a very large percentage of feebleminded children, a much larger percentage than the average of other occupational groups. But it must be remembered that the intelligence tests used do depend upon efficiency in language for their scores and the children of these two groups are most likely to be weak in this respect.

If all the occupational groups represented in this study are classified into the five groups described by Taussig[1] we obtain Table III. It will be noted that up to Class IV there is an increase in IQ in proportion to the increasing demands for skill and planning in the occupational levels. Comparable findings were obtained from Stuart M. Stoke's[2] investigations of the children who lived in a suburb of Boston. The two sets of results have been placed side by side in Table IV. There are certain significant similarities. In the lowest occupational level, Group I, no child appears in the very superior group from either investigation; in the superior group 5 per cent from the present study occur and none from that of Stoke. The correspondence in Group I, considering the rough way in which the classifications are made is markedly similar.

In the semi-skilled group (Group II) there are also close similarities in the highest levels. The percentages of very superior and of superior are as nearly alike as could be expected. The marked differences between these two investigations in Group II is in the borderline group.

[1] Taussig, F. W., Principles of Economics. Vol. 2, pp. 134-137.

[2] Stoke, Stuart M., Occupational Groups and Child Development. pp. 14-21, 1927.

TABLE III

Distribution of 1247 Children According to Intelligence Quotient Group and Occupational Group

(Taussig's groups plus the farmers, No. VI.)

IQ	I	II	III	IV	V	VI
Above 119		17	11	16	13	2
Very Superior						
110-119	1	54	19	20	25	11
Superior						
90-109	9	261	106	60	53	46
Average						
80-89	6	149	47	20	18	26
Dull						
Below 80	4	154	38	10	16	35
Borderline						
Deficiency						
Number	20	635	221	126	125	120
Median	90	91	94	103	100	90

TABLE IV

Distribution of 1247 Children According to Intelligence Quotient Group and Occupational Group. Percentages.

(Taussig's groups and Farm Group, No. VI)

IQ	I	II	III	IV	V	VI
Above 119	0	2.67	4.97	12.67	10.4	1.67
(Very Superior)	0	2.7	5.0	6.1	28.0[1]	
110-119	5	8.48	8.59	15.86	20.0	9.16
(Superior)	0	8.2	14.5	16.9	32.0	
90-109	45	40.98	47.91	47.58	42.4	38.32
(Average)	57.9	55.5	62.9	66.7	36.0	
80-89	30	23.39	21.24	15.86	14.4	21.66
(Dull)	26.3	28.1	13.8	7.2	4.0	
Below 80	20	24.18	17.17	7.93	12.8	29.16
(Borderline Deficiency)	15.8	5.5	3.8	3.1	0.0	

[1] Figures in second row each time from Stoke, Stuart M. Occupational Groups and Child Development. p 16, 1927.

In the present investigation the percentage of Group II falling in this category is 24.18 while in Stoke's investigation this percentage is 5.5. In Groups III, IV, and V the outstanding difference between the two studies appears in the borderline cases which are far more numerous in the present study. Group IV in both cases are somewhat similar with a larger percentage of the former investigation appearing in the very superior group and a smaller number in the borderline group. In Group V, the business and professional group, do we find the most significant differences. There are evidently very great differences between small town storekeepers, ministers, and physicians of the present investigation and those professional and executive groups residing in areas suburban to Boston. For example, the present investigation finds 27 per cent of children of parents classified in Group V among the dull and borderline while the former investigation finds only four per cent among the dull and none among the borderline.

Some insight into the cause of the very low scores earned by these children may be had by separating the IQ's into those earned on (a) language tests and on (b) non-language tests. Such a procedure has been carried through in Table V.

In this table about the same theoretical expectancy is used as Woodrow used in "Brightness and Dullness in Children." In the first column we have the theoretical expectancy in percentages. In the second column appear the percentages calculated from the scores of Pintner-Cunningham and Dearborn; in the third column the results from a test involving written language (National Intelligence Test); and in the fourth the results of a language test (National Intelligence Test) given in another community by another investigator (Jobe, 1922). In comparing the non-language results with those of theoretical expectancy, children of mill workers have a slightly larger percentage of very superior (2.4 *vs.* 6.8),

somewhat larger per cent of superior, a smaller percentage of normal, and a larger percentage of dull, borderline, and feebleminded. If we compare the language tests with theoretical cases we find both in 1922 and

TABLE V

Percentage of Children of Mill Workers at Each Intelligence Level

Year		1925 N103	1925 N243	1922 (Jobe)[1]	
IQ	Expectancy	Non-Language	Language	Language	Average Last Two
Genius (Above 139)	.1	0	1.2	0	.6
Very Superior (120-139)	2.4	6.8	.4	2.2	1.3
Superior (110-119)	15.0	16.5	5.3	3.7	4.5
Normal (90-109)	60.0	47.6	42.3	41.5	41.9
Dull (80-89)	15.0	16.5	22.6	20.6	21.6
Borderline (70-79)	5.0	8.7	16.0	19.5	17.8
Feebleminded (Below 70)	2.5	3.9	11.9	12.5	12.2

1925 a smaller percentage of very superior, superior, and normal; a larger percentage of dull, a much larger percentage of borderline (5 *vs.* 16 or 19) and a much larger percentage of feebleminded (2.5 *vs.* about 12). In proportion as the tests depend upon language the children of mill villages suffer in comparison with theoretical expectancy. Of more than ordinary interest is the close agreement between the levels of intelligence in two studies undertaken about three years apart in mill villages quite differently located and involving in each case more than 500 children. One could hardly expect two sets of data to agree more perfectly (columns 3 and 4).

[1] Master's essay at the University of North Carolina.

In addition to the evidence just now presented relative to the influence of language upon intelligence scores further investigations have been made. Among our data there were found a considerable number of cases from this community who had taken both the tests involving written language (National Intelligence Test) and those not involving it (Pintner-Cunningham and Dearborn). We have thus displayed side by side intelligence quotient scores obtained from language tests and from non-language tests.

In Table VI are found the IQ scores of 181 children who had been tested on the Dearborn Test, Series I, a non-language test, and upon National A, a test which requires the mastery of simple language habits in order to make a good score. It will be observed that the median score of the Dearborn test is 98.6 and of the National A, 88.2, a difference of more than ten points. If these differences were between individuals, the ten points would not be so striking, but such a difference between averages into which 181 scores had entered becomes immediately significant. While the difference between the means is not quite so large (8.75) it is a significant difference (σ dif $= \pm 1.43$).

TABLE VI

Differences in IQ's of the Same Children on (1) Language Tests and, (2) Non-Language Tests

	Mean	N	Dif. Means	P.E. Dif.
Dearborn (non-language)	98.05	181		
National A (language)	89.3	181	8.75	±1.43
Pintner-Cunningham (non-language)	98.3	227		
National A (language)	87.7	227	10.6	±1.36
Pintner-Cunningham (non-language)	99.08	288		
Dearborn (non-language)	97.3	288	1.78	±1.08

Even more substantial is the difference between IQ scores on the Pintner-Cunningham and the National A.

(Table VI). In this instance the difference between averages is 10.6 IQ points with a standard error of 1.36. The former figure is about eight times the size of the latter which indicates a reliable difference between IQ scores of individuals taking both the Pintner-Cunningham and National A.

In Table VI are also displayed 288 cases in which the pupils received IQ scores on the two non-language tests. Here we find instead of substantial difference, substantial agreement. Instead of a difference of 8 or 10 points between the averages we have a difference of less than two points, together with a standard error almost as large. The difference between the averages in this instance is not twice the size of its standard error and hence is *not* reliable. In short these two tests are in substantial agreement as far as the mean or median performance is concerned.

From the evidence thus far presented it seems clear that the children of our investigation score somewhat lower on the language tests that were used than upon the non-language tests. This might be due to one of two things: (a) the language tests might have higher standards or (b) because of an inferior environment particularly in the way of reading facilities the pupils score less well on the language tests. In the present instance the second factor is undoubtedly the dominating one for the children included in our population were from rather unstimulating environments. More than sixty per cent of the parents of these children lived in homes not their own and received salaries barely adequate to furnish the essentials of life. The parents, too, in many cases had failed to secure an adequate education so that the only intellectual stimulation received by the children must needs come from the school. The fact also that two tests (Dearborn and Pintner-Cunningham) which do not depend upon written language agree so closely is additional evidence of this con-

tention. A good illustration of the difference in scores appears when the number of feebleminded (IQ of 70 or below[1]) is considered. On the Dearborn test out of 181 children only one is below an IQ of 70; on the National, 16 are below an IQ of 70. On the Pintner-Cunningham test out of 227, six are below an IQ of 70; while on the National A, 27 are below this point.

For these reasons it is clear that the interpretation of an intelligence test must be in terms both of the type of test used and of the environment of the pupils studied. It seems evident that a non-language test should also be used in every case where the environment is even suspected of being below that on which the tests were standardized or else such absurd differences occur as appeared in the above paragraph.

Finally we see a combination of language and other environmental factors at work in Table VII. In these tables the median IQ's of each age for both the mill and the non-mill are presented.

TABLE VII

Decrease of IQ's with Age

Age	6	7	8	9	10	11	12	13
Number	41	77	71	65	53	71	62	54
Mill (Md)	100	99	95	91	85	90	87	85
Number	70	123	117	107	108	95	101	90
Non-Mill (Md)	101	103	100	94	92	94	94	92

The children of textile workers show a decreasing IQ from year to year. If we consider the years from 6-13, ages in which the number of pupils represented is of substantial number, the median IQ drops from about 100 to 85. According to this table, in the ordinary environment the median child at year 6 would be amply able to hold his own in any school system in

[1] The argument would be the same if the IQ for feeblemindedness were placed at 60 IQ.

the country, but by the time this same child had reached the 13th year he would be definitely retarded and regarded as a dull child able only with assistance to carry on the work of the school. In part, to be sure, the difference between IQ's is accounted for by a change in tests after the second year and the sharp drop may be attributed to the language element appearing but even this language influence might be one of the clearest signs of an impoverished cultural background. With the non-mill group there are small decreases up to age 13 until the IQ that started at 100 is then only 92. The drop in IQ points from years 6-13 is in the case of children of mill workers 14.7 points; in the case of non-mill workers, 8.6 points. Two hypotheses might be offered in explanation of these rather curious phenomena. On the one hand, we might attribute the differences to a poverty stricken environment. Workers in cotton mills get very meager wages and these bring in their wake environments lacking in many cultural contacts found in other environments. This environmental lack is probably increased in the majority of cases by a failure on the part of the employees to own their homes. Then, too, many move around a very great deal and hence accumulate the minimum of personal property. These conditions are certainly not conducive to the development of a desirable sort of cultural contacts. Now these conditions prevent the children from developing as rapidly in reading, arithmetic, etc., as they would in some other communities and hence when a language test depending on a knowledge of reading and arithmetic is given, these children fall behind in a deplorable manner. The second hypothesis is that there really is a much larger number of feebleminded in the upper grades since the tests are very little dependent upon environmental factors. This hypothesis certainly does not ring true for how could children who were almost normal at seven years grow into feebleminded ones at 10 or 11? If there

is one measure more constant than most others concerned with human beings that measure is the IQ. Furthermore, these language tests indicate a far larger per cent of definitely feebleminded children than common sense would indicate, for certainly there is no observable indication in mill children that as many as 12 per cent are feebleminded. These findings then are preposterous on their face and such an hypothesis is unsubstantiated. The first hypothesis seems to be the more plausible one.

Summary and Conclusion

1. More than 1200 school children in Grades I to VII have been tested with standardized tests of intelligence. These records have been tabulated and inspected for the purpose of discovering the influence of parental occupations on the scores of children.

2. Substantial differences in median scores appeared among the various occupations. For example, the children of the professional group scored 105; while those of mill workers and farmers scored 91 and 88 respectively. Again the children of travelling salesmen, insurance agents, merchants, bankers, and lumbermen, plumbers, and machinists scored in the neighborhood of 100; while those of carpenters, clerks, stone cutters, blacksmiths, and laborers scored in the neighborhood of 90.

Many of the published comparisons between children whose parents differ in occupations have been made between rural and urban children. On tests involving language there is always an average difference in the neighborhood of 10 IQ points which as you remember is the same as the difference between the scores on the language and on the non-language tests. On non-language tests the differences are not so large but are still in favor of the urban group.[1]

[1] For further references, see:

1. Jones, E., Conrad, H. S., Blanchard, M. S., "Environmental

3. When the children's scores were classified into IQ's above 139 (Genius), 120-139 (very superior), 110-119 (superior), 90-109 (normal), 80-89 (dull), 70-79 (borderline) and below 70 (feebleminded) further important differences between the occupations were shown. Among the feebleminded group the farmers' and mill workers' children showed the largest percentages, 15 and 9.5 respectively. In the genius group only the children of the professional workers were substantially represented.

In general, our whole group shows more among the three lowest groups and the two highest groups but fewer among the dull and average groups. The children of mill workers so abundantly represented in this population show a larger percentage than expectancy among the dull, borderland, and feebleminded (nearly 48 per cent instead of 22.5); a smaller percentage than expectancy among the normal and superior (nearly 50 per cent instead of 75) and a slightly higher per cent in the very superior and genius groups (2.7 instead of 2.5 expected).

4. When the scores of the children were classified into groups according to the economic level of their parents (Taussig's classification) it was clear that in general the median scores increased *pari passu* with economic level. The median scores of farmers' children

Handicap in Mental Test Performance." University of California Publications in Psychology, Vol. 5, pp. 63-99, 1932.

2. Pintner, R., Intelligence Testing, pp. 251-255, 1931.

3. Hinds, J. H., "A Comparison of the Brightness of Country and City High School Children," Jour. Ed. Research 1922, *5*, 120-124.

4. Pressey, S. L. and Thomas, J. B., "A Study of Country Children in (1) A good and (2) a poor farming district by means of a Group Scale of Intelligence," Jour. Applied Psy. 1919, *3*, 283-286.

5. Shimberg, M. E., "An Investigation of the Validity of Norms with Special Reference to Urban and Rural Groups," Archives of Psychology, 1929, No. 104.

approached most closely that of Taussig's Unskilled Group (I). Comparison was made with the findings of Stoke in a New England community by computing and displaying the percentages appearing in each IQ level. At Taussig's Level I (unskilled) there was no great difference. In Level II (semi-skilled), our most populous group composed so largely of mill workers, the percentages at each level were closely similar except that the present study showed a much larger percentage in the borderline deficiency group. The greatest differences appeared among the professional and manager groups of the two studies. The New England study showed a definitely larger per cent of membership in the superior groups and a very much smaller percentage among the dull and borderline groups.

5. In the present study a much larger percentage of children fall in the dull and feebleminded group than in the case of theoretical expectancy. Three lines of evidence which converge upon the same focus have been submitted in explanation of these phenomena.

(a) When we consider the median scores of children of mill workers, for example, we find that they decrease from about normal IQ's at year 6 to 85 at year 13.

(b) When the same children are tested with both language tests and non-language tests there are from 9-10 IQ points difference in favor of the non-language tests.

(c) When comparisons were made between language and non-language tests there were shown to be an impossible number of feebleminded present among the children of mill workers in the former test and a more reasonable number in the latter.

6. These three lines of evidence lead us to the conclusion that in those occupations represented largely by semi-skilled workers the poverty of the environment is so severe that children's scores on intelligence tests are definitely lowered so that children who begin their work in school with normal capacities have dropped to the level of dullards by the time they have attained the age of thirteen. This means that the tests of intelligence which depend on written language have been standardized at the several ages on children whose environment was definitely superior to that represented in the present study. The effect of this poor environment is accumulative and causes smaller IQ's in successive years.

RESULTS OF CERTAIN HONESTY TESTS GIVEN TO A GROUP OF RURAL WHITE CHILDREN AND TO TWO GROUPS OF INDIAN CHILDREN

S. L. and L. C. PRESSEY
Ohio State University

INTRODUCTION

During the year 1932 the writers were carrying on an investigation into the emotional attitudes and interests of both Indian and white children.[1] The main part of this research was based upon results obtained by use of the Interest-Attitude Tests (a greatly improved form of the X-O Examination published some years ago.[2]) However, since this test being most widely used was verbal in character, it seemed desirable to make some further study of Indians and whites by means of performance tests. Because the May-Hartshorne[3] materials dealing with honesty were ready for use and easy to administer, it was decided to use three of these tests as part of a preliminary investigation into any differences of this character that might occur. The three tests were selected so as to cover different situations in which honesty was required, namely: (1) honesty in scoring one's own arithmetic test, (2) honesty in handling money, and (3) honesty in keeping one's eyes closed when one has been instructed to do so.[4]

[1] Under grant from the National Research Council.

[2] See Pressey, S. L. and L. C. "The Interest-Attitude Tests". Journal of Applied Psychology, Vol. 17, pp. 1-16, Feb. 1933.

[3] Mays, M. A. and Hartshorne, H. *Studies in Deceit, Book I.* New York Macmillan, 1928, pp. 414.

[4] A fourth test, involving individual testing with the dynamometer was tried, but was given to only 151 whites and 55 Indians. The writers wish to express their obligations to Dr. D. A. Worcester of the University of Nebraska for his cooperation in this work.

These three tests were given to one group of 194 rural white children attending the third, fourth, and fifth grades in five small towns in Nebraska. They were given also to two groups of Indian children in the same grades in school. One group consisted of children tested in four different schools in Nebraska. There were 208 cases, the great majority of whom (96 per cent) were Sioux, Omahas, Winnebagos or Pottowatomies. The second group of Indians consisted of 197 Navajos tested in two schools in New Mexico.

The results will be presented by test and will deal with (a) changes in honesty from grade to grade for all three groups, (b) comparison of the whites and Indians in Nebraska, and (c) comparison of the two groups of Indians with each other. Certain general conclusions regarding honesty in the three situations presented by the tests will also be given.

RESULTS

Test I: *Honesty in scoring one's paper.* In administering this test the children were first given two sets of addition combinations, one of which presented the figures in exactly reverse position from the other; the larger figure was above the smaller in the same number of instances on both sets. The children were given one minute on each form, after which the sheets were collected. These two forms were given under rigid supervision. About 15 minutes later the children were given a third set of combinations which consisted of items chosen equally from the two earlier forms; the same time limit was used, but the children were then given answer sheets and allowed to score their own papers. The supervision was sufficient only to prevent conversation but was lax enough to allow cheating. The score on this first test consists merely of the *difference* between the score on the test the child corrected and his score on the *higher* of the two forms known to be taken under "honest"

conditions. A distribution of these differences is presented below for each grade and each group.

TABLE I

	Grade 3			Grade 4			Grade 5		
	White	Sioux*	Navajo	White	Sioux	Navajo	White	Sioux	Navajo
36			1						
30-35			2						
25-29			9						
20-24		1	2			1			2
18-19			6			3			1
16-17			5	1		3			1
14-15		1	2		1	4		2	1
12-13	1	2	4			3	1	2	4
10-11	1	2	7		1	3	1	1	7
9		2	4		1	5			2
8	1	6	5		4	1	1	3	3
7	3	2	3	3	2	3	2		1
6	4	4	10	1	4	4		3	4
5	4	7	4	3	7	2	1	5	4
4	9	7	4	2	6	4	2	5	5
3	7	7	2	8	4	2	2	4	5
2	5	8	1	10	6	2	4	5	6
1	9	9	2	5	6	4	8	5	2
0	7	5	5	14	3	1	16	8	3
–1	4	2	2	3	2	1	7	4	1
–2	3	1	3	3	4	2	3	3	
–3—4	3	2	2	1	3		5	1	
–5—6	1	3	1		2		2	2	
–7—8	1	2		2	2		1		1
Totals	**63	73	86	56	58	48	56	53	53
No. Dishonest	14	27	64	8	20	32	6	16	30
% Dishonest	22	37	75	14	34	67	11	30	57

* Under "Sioux" are included all results from the Nebraska tribes.

** These totals do not agree with those given for the other two tests. Each arithmetic test used contained 36 items. Some chil-

Those children who increased their better previous score by 5 points or more were regarded as dishonest. A line has therefore been drawn across the table just above the "4" interval. It will be seen that for all three groups there is a decrease in the percent of dishonest scores from grade to grade.

In the third grade, 15 per cent more of the Nebraska Indians were dishonest than the rural Nebraska whites; in the fourth grade, 20 per cent more, and in the fifth grade, 19 per cent more. It should be kept in mind throughout the comparisons between these two groups that the Indians averaged 2.1 to 2.3 years older than the whites and would, were the two groups equally honest, score distinctly lower because of their added years. In other words, the average difference of nearly 18 per cent is not a true picture of the real difference; if the two groups of children had been of the same chronological age the difference would undoubtedly have been greater.

Finally, a comparison of the Nebraska Indians and the Navajos reveals interesting differences. Almost exactly twice as many Navajos (67 per cent of the total in all three grades) cheated than did members of the other group (34 per cent of the total). The writers would like to suggest that the Nebraska Indians live in relatively small groups surrounded by and intermingling with whites. The children sometimes go to the same schools as the whites; they often play with white children. They are almost inevitably taught by white teachers. Both the children and their parents are constantly subject to the pressure of "white" standards. As a result, the Indians gradually conform, but the process

dren made scores of 32 or more on at least one of the forms taken under "honest" conditions. For these children there was no opportunity for cheating to show itself, since the greatest possible increase would be 4, 3, 2, 1, or 0 points. These cases have, therefore, been omitted inasmuch as nothing is known concerning their honesty.

takes longer. It is likely that results from the eighth grade, for instance, might show little difference between the white and Indian children. On the other hand the Navajos have an isolation and a tribal solidarity not approached by any of the groups in Nebraska. Book-learning, arithmetic, tests, scores on tests—these are all matters of their schooling independent of their home training. In these early grades their response to school training is relatively slight. As a result of the differences in the two social situations in which these two groups of Indian children have grown up, one finds greater differences between the two Indian groups than between one of these groups and the neighboring whites. One must certainly conclude that Indians are not "naturally" dishonest, but that their honesty in this or that situation will vary from the norms set by white children in proportion as environmental influences have been controlled by "white" standards of conduct.

Test II: *Honesty in handling money.* The test used for determining honesty in this situation was, in appearance, another test of arithmetic. The children were given sheets on which appeared a dozen such questions as: "Which three pieces of money make exactly 16c?" Each sheet had a code number written in one corner on the back. Each child was given a small pill box containing one quarter, two dimes, two nickels, and two pennies. The box had a sliding cover on the inside of which, in a position invisible to the user, was written the same code number as appeared on the sheet given each pupil. Greatest care was exercised to get these sheets and boxes correctly distributed; in no case was there any need for the assistants to examine the code numbers before the children in order to have papers and boxes match. When the children had their materials they counted out of the box the dime, nickel, and penny needed to make exactly 16 cents; then they wrote down the answer to the first question. A second question was similarly answered.

Then the children were instructed to answer the remaining questions by counting out the necessary pieces of money and then writing the answers in the spaces provided on the sheet. The score on this test is merely the amount of money stolen; the answers to the questions were irrelevant and were never scored. The amount of money stolen by each child in each of the three groups is shown below:

TABLE II

	Grade 3			Grade 4			Grade 5		
	White	Sioux	Navajo	White	Sioux	Navajo	White	Sioux	Navajo
50c			1						
40c			3						
30c			3						1
25c		1	1		1	5		2	4
20c									
15c			2						
10c	1	5	7		2			2	5
5c			5	1	2	1			
0c	66	68	67	63	60	51	63	65	54
Total	67	74	89	64	65	57	63	69	64
No. Dishonest	1	6	22	1	5	6	0	4	10
% Dishonest	1.5	8	25	1.5	8	11	0	6	16

The first point is again a general tendency for dishonesty to decrease with age in each group. The Nebraska Indians are more dishonest than the whites, but not nearly as dishonest as the Navajos. Presumably the same explanations hold as given for the earlier test. For in no group is the amount of dishonesty as high as for the other tests.

Test III: *Honesty in keeping one's eyes closed when instructed to do so.* The test used consisted of a single sheet of paper on which seven squares were printed, one

inside the other. The children first drew around the innermost square with their eyes open, trying not to touch the sides of either the inscribed or circumscribed square. They next put their pencils at the starting point for drawing a line between the second and third squares, but were told to keep their left hand over their eyes while they did the drawing. Each successive square was drawn around in the same way—locating the starting point with eyes open and then *supposedly* drawing with the eyes closed. The examiner paid no attention to any cheating, no matter how blatant it was. The score on this test consists merely of the number of corners where the child's pencilled line does not touch the printed enclosing lines on either side. Since there were five test squares, scores could run from 0 to 20. The distribution of scores is given below. It had been found from previous testing that a score of more than 4 indicated dishonesty, the degree increasing as the number of correctly drawn corners becomes larger.

The results from this test show rather less decrease with age than those from the two previous tests. The writers have an idea that the explanation lies in the novelty of the situation. Children are specifically and thoroughly trained by home, school and community to be honest in handling money; the school has to struggle alone with dishonesty in scoring tests, but the efforts at securing honesty in this respect are intensive; nowhere, however, except in games is there specific training in not peeping through one's fingers. Honesty in this situation, especially when presented in school, is the result of transfer—or else of a general "concept of honesty." Throughout the years of childhood there is a decided specificity in moral reactions. Many children did not perhaps recognize the "honesty" element in this new situation and many who did could not make adequate transfer. In general, then, one might deduce that the more specific the training the greater the honesty in any particular

TABLE III

	Grade 3			Grade 4			Grade 5		
	White	Sioux	Navajo	White	Sioux	Navajo	White	Sioux	Navajo
20	9	11	68	2	10	23	2	4	33
19	1	2				2	1	1	3
18		4	2	1	4	2		2	3
17	1	2	1	2	1		2	2	3
16		1	1		2	2	2		3
15		2	2		2		2	1	1
14	1	2			2	2	3	2	2
13	1	2	1	1	1	1		1	
12	1	3			1	1	2	1	1
11	1	2	2	3	3	1		2	
10	2	3	1	3	1	3		1	3
9	2	2	1	2	2	3	1	3	1
8	1	5	2	3	1	1	2	3	
7	1	3	2	1	3	2	1	3	1
6	1	3	2	2	2	2	2	4	1
5	2	3		2	1	1		6	
4	1	5	1	3	4	3	2	8	2
3	3	6		8	5	2	5	13	
2	13	5	1	9	4	2	5	6	
1	14	3		4	12	1	4	3	2
0	12	5	2	18	4	3	27	3	5
Total No.	67	74	89	64	65	57	63	69	64
Dishonest	24	50	85	22	36	46	20	36	55
% Dishonest	36	68	96	34	55	81	32	52	86

situation. This not only seems reasonable but is borne out by such figures as those from the white, fourth-grade children of whom 1.5 per cent stole money, 14 per cent cheated in scoring their own papers, and 34 per cent cheated by peeping through their fingers. Other grades, from all groups, show a similar tendency.

This test shows the Nebraska Indians distinctly less honest than the whites. Presumably, the same explanation holds as was just given above—the unfamiliarity of

the situation and the inability of the children to make a transfer from more usual situations involving honesty. In this connection, the greater chronological age of the Indians should again be remembered. However, for sheer, complete, and universal dishonesty the Navajos are outstanding. Their carefully acquired honesty in money matters or in other specific situations and their honesty in their own tribal relationships simply went to pieces before this new test. The amount of dishonesty runs from 96 per cent of the children to 81 per cent. Out of the 197 scores, 59 per cent were "20"—that is, all five squares were perfectly drawn—indicating that the children were looking at their work continually, another 30 per cent were more or less dishonest, while only 11 per cent from all grades combined show honest scores.*

CONCLUSIONS

Obviously, more work needs to be done before any final decision may be arrived at. However, the authors would like to offer the following tentative conclusions:

(1) Children of both races are more honest in tests for which their training has fitted them than in tests for which they have received no specific training.

(2) For both races and on all tests, dishonesty decreases with age, but to a smaller extent if the "honesty" element in the test situation is not obvious.

(3) Indian children living among whites gradually adopt "white" standards, but not until they are chronologically older than the white children manifesting these standards.

(4) Indians living under conditions totally unlike those of white children show highly different scores on

* The fourth test, mentioned earlier, showed 13 per cent of the 151 white children to have been dishonest when allowed to record their own results with the hand dynamometer, as compared with 33 per cent of the 55 Indians.

these tests. In fact, the three groups of children tested vary from each other in proportion as their environment is dissimilar.

(5) The Indians from Nebraska are more like the white children of Nebraska than they are like the Navajos.

(6) The results accord with observation. Rural white children in the Middle West are usually honest. Navajos are notoriously dishonest.

(7) There is no evidence that honesty is an inborn trait or a racial characteristic. There is, on the contrary, strong evidence that a child is honest in proportion (a) as he has been trained to honesty, and (b) as he is able to transfer his training to the particular situation before him.

(8) It is likely, in the writers' opinion, that Indian children at the end of their school will show more dishonesty in such tests as have been described than white children because they begin school at a later age and because they have to "straddle" two cultures that are often divergent on the particular questions of honesty involved.

(9) This opinion is not, however, to be understood as bearing in the least upon any racial difference. If the Indian pattern of behavior were the norm, the white child in similarly difficult surroundings would presumably conform only slowly and imperfectly.

THE EFFECTS OF BUSINESS CONDITIONS UPON THE SANITY OF A POPULATION

FREDERICK J. GAUDET
Dana College, Newark, N. J.
and
MARCUS A. CURRY
State Hospital, Greystone Park, N. J.

The question of the effects of business conditions upon the sanity or insanity of a civilized group has been discussed by many authorities. If one looks through the annual reports of the superintendents of state institutions for the insane, one frequently finds that an increase in the hospital population, or in the admission rate, is said to be due to a period of business depression. The evidence for such a statement is always the present condition—a single instance. So far as the authors are aware, no one has attempted to gather any evidence pointing to the fact that business depressions invariably bring about a greater hospitalization.

In support of this view that business depressions bring about a greater amount of insanity in our various state or hospital districts, clinicians add their evidence. They point out that a great many of their patients place the cause of their breakdown upon the stress which they underwent due to a fall in the market, loss of a job, part-time work, etc. However, others have pointed out that it has been only a few decades since clinicians listened to stories of their patients regarding overwork as a cause of a breakdown of mental balance.

In order to get evidence of a more reliable nature, the authors have attempted to find what relation there is between business conditions and hospitalizations over a period of years. The data used in this study have been

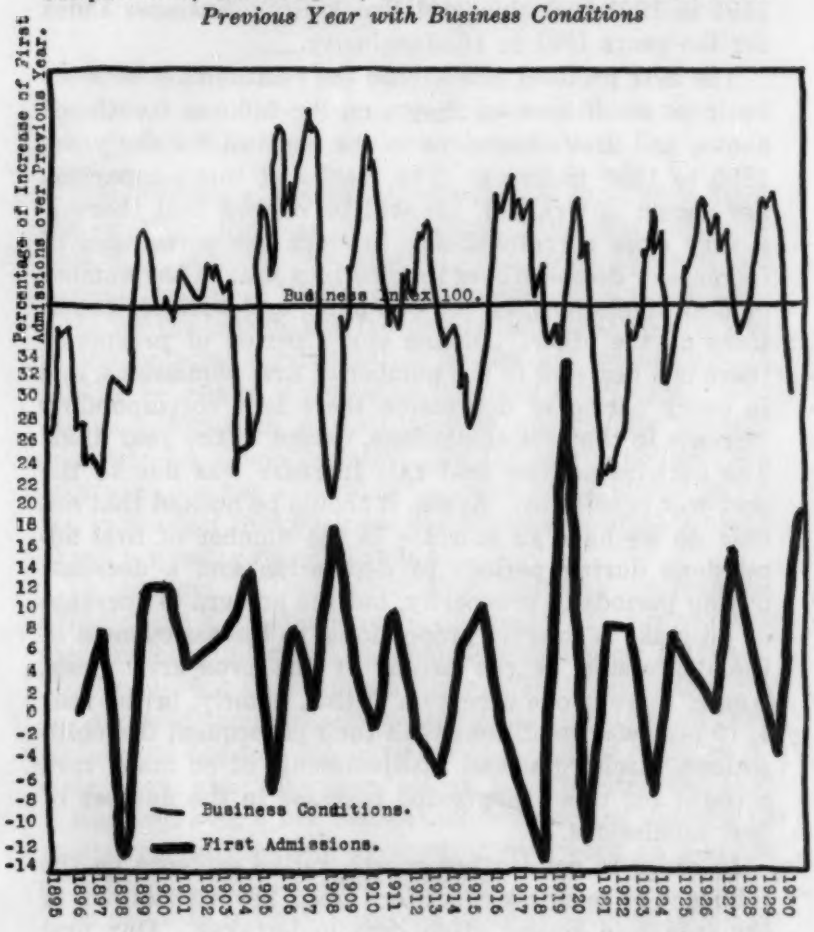
taken from the annual reports of the New Jersey State Hospital at Greystone Park, New Jersey, for the years 1895 to 1930 inclusive. As a measure of business conditions, Snyder's Clearing Index has been used for the years 1895 to 1901 inclusive, and Brookmire's Business Index for the years 1901 to 1930 inclusive.

The first problem was to find the relationship between business conditions, as shown on the indexes mentioned above, and first admissions to the hospital for the years 1895 to 1930 inclusive. The results of this comparison are shown in Graph I. It will be noticed that there is a very close correspondence between the percentage of increase or decrease over the previous year in the number of first admissions, on the one hand, and business conditions on the other. During every period of prosperity there is a decrease in the number of first admissions, and in every period of depression there is a corresponding increase in the first admissions, except in the year 1899. The authors assume that this increase was due to the post-war conditions. Again, it should be noticed that not only do we have an increase in the number of first admissions during periods of depression and a decrease during periods of prosperity, but the amount of increase or decrease is roughly proportional to the seriousness of the depression or the height of the prosperity peak. Again, there is one exception to this, namely, in the year 1919 post-war conditions with their consequent demobilizations, discharges, and readjustments of so many men account for this exaggerated increase in the number of first admissions.

In order to get further corroborative evidence on the influence of business conditions upon hospitalizations of the insane, a second study was undertaken. Our first study upon the relationship between the number of first admissions and business conditions points out that hospitalizations increase, but does not offer any evidence to prove that business depressions can act as etiological

GRAPH 1

Comparison of Percentage of Increase in First Admissions over Previous Year with Business Conditions



factors. This increase, which is found during business depressions, might be explained as being due to the poorer financial status of the families during such periods, with the result that a larger number are hospitalized who, under more normal business conditions, might be kept at home. The authors do believe that this may account for a slight increase in the number of admissions, and they also believe that a number of those admitted would be sent to private institutions in periods of better business conditions, but do not believe that this accounts for the very great increases that accompany such periods.

In order to get further information upon the etiological effect of business conditions, the authors studied the relationship between business conditions (as shown in the indexes) and the median age of patients when admitted. The results of this comparison are shown in Graph II.

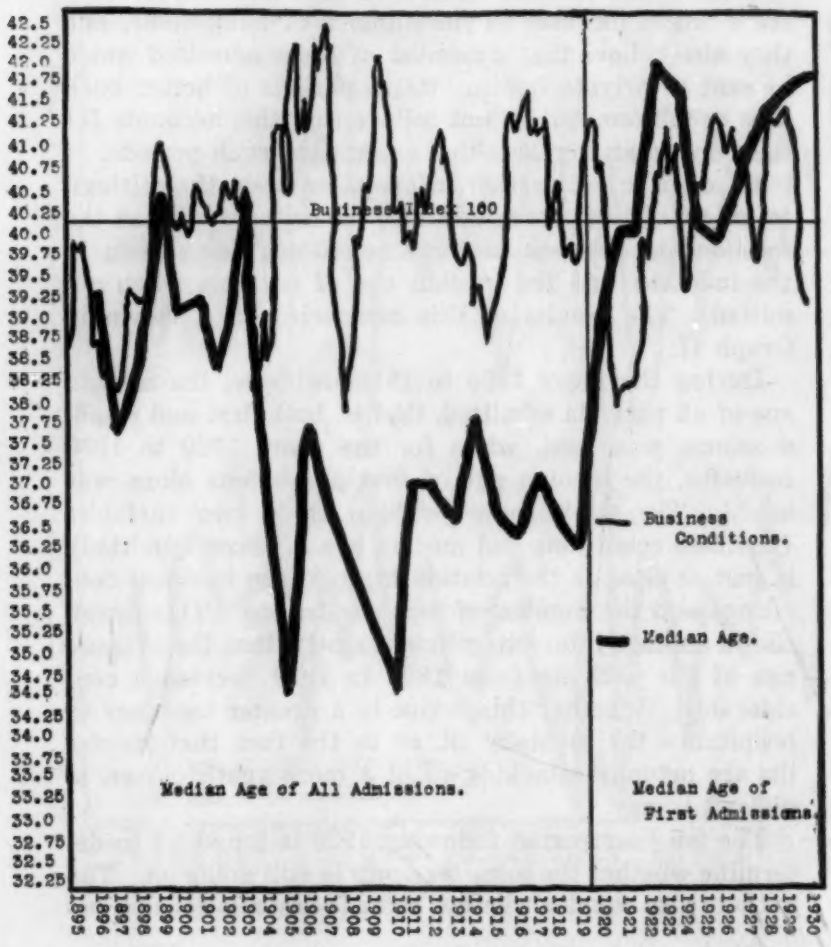
During the years 1895 to 1919 inclusive, the median age of all patients admitted, that is, both first and re-admissions, was used, while for the years 1920 to 1930 inclusive, the median age of first admissions alone was used.[1] The relationship between these two variables (business conditions and median age of those admitted) is just as close as the relationship between business conditions and the number of first admissions. This graph shows another interesting item, namely, that the average age of the patients from 1895 to 1919 decreased considerably. Whether this is due to a greater tendency to hospitalize the mentally ill, or to the fact that mental ills are actually attacking us at a more youthful age, is difficult to say.

The ten-year period following 1920 is too short to determine whether the same tendency is still going on. The authors also wonder whether there is any relationship be-

[1] The reason for this was that the reports prior to 1920 gave age distributions of all admissions and after 1920 age distributions for first admissions only.

GRAPH 2

Comparison of Median Age of Patients Admitted to Greystone Park and Business Conditions



tween the decreasing age of the first admissions to hospitals for the mentally ill and the decreasing age of apprehended criminals.

Further studies were made in attempting to find the relationship between business conditions and the number of readmissions, but since there is another factor, namely, the number of discharges in the previous few years, which influences this, the results were of little use. This study did bring out the fact which every clinician and hospital superintendent knows, namely, that discharges decrease during times of business depression. It was also noticed that the number of re-admissions depended to a great extent upon the number of discharges in the preceding year.

In conclusion, the authors wish to point out that they do not believe that these two studies *prove* any relationship between business conditions and the sanity or insanity of a population. They do believe, however, that this research adds further evidence to the previous *opinions* of superintendents and clinicians. Similar data obtained from other institutions would add considerably to such evidence as has been presented here. A study of admissions to all hospitals, both private and public, would add more value to such a study, which would be able to show whether the increased hospitalizations found in periods of depression are due, to some extent, to the fact that fewer people go to private institutions during periods of business depression.

FOUR NEW TESTS OF MUSICAL TALENT

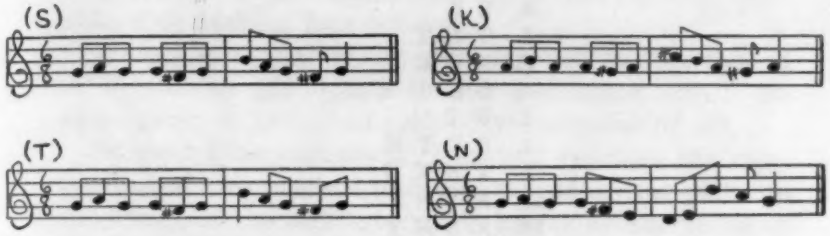
RALEIGH M. DRAKE
Wesleyan College

INTRODUCTION

The four tests described here are attempts to develop some reliable measure of musical talent. Each test represents a different approach to the problem in the hope that some cue to the essential nature of musicality will be disclosed. All have been carefully studied with respect to validity and reliability and compared to other prominent tests already published by other investigators. An account of this comparison will appear in a forthcoming article and the purpose of this paper is to describe the tests.

DESCRIPTION OF TESTS

Musical Memory. This test follows the practice that is considered good testing procedure of having the material as near like the function to be tested as possible. The test consists of twenty-four trials for which there are as many original two-bar melodies. All of the melodies were originally composed to obviate any possibility of familiarity with the material. Each one of these twenty-four melodies serve as a standard of comparison with four possible variables: change of key; change of time; change of note; or the same melody repeated. Thus in trial 9 the original melody or standard is:

to which each of the following are to be compared and identified (the original melody is not repeated for each comparison):

The first variable is the *same* melody repeated, the second is the same melody played in a different *key*, the third variable has the same notes as the standard but the *time* values of some of the notes are altered, and the fourth variable has five *notes* different from the standard.

Each variable is identified by the first letter of the four possible variables. Thus, S stands for Same as the standard, K stands for change of key, T stands for change of time, and N stands for a change of one or more notes. Only one kind of change is made in any one variable. All that need be recorded for trial 9 would be S, K, T, N, which are placed in a specially prepared form as shown in Fig. 1.

Figure 1 shows the correct responses to all 24 trials, illustrates how serial order is worked out, and indicates how with each four trials there is one additional variable required to be identified. The first four trials each have only two variables following the standard and each variable comes both immediately after the standard and second from the standard. The next four trials have each three variables following the standard, each variable coming immediately after, second, and third, following their respective standards. This increasing number after each fourth trial continues together with the serial order

FIGURE 1

1. K S
2. N T
3. T N
4. S K
5. T N T
6. K S K
7. S K S
8. N T N
9. S K T N
10. T N K S
11. N T S K
12. K S N T
13. N T K S K
14. S K N T N
15. K S T N T
16. T N S K S
17. K S T N K S
18. N T K S N T
19. T N S K T N
20. S K N T S K
21. T N K S T N T
22. K S N T K S K
23. S K T N S K S
24. N T S K N T N

until there is a maximum of seven variables following the last four trials, thus making the test increasingly difficult.

All standards are only two-bars in length and increase slightly in difficulty or complexity, from one trial to the next. When there are more than four variables following a standard, as in trials 13 to 24, some of the four variables have to be repeated (which ones are repeated depends on the serial order that works out for that trial), so the changing variables (K, N, T) are altered in a different fashion so that no two variables of the order K, N, or T are ever exactly alike.

The test is given on a piano. It is a group test and can be given to 75 or 100 subjects easily if acoustic conditions are favorable. The time necessary to explain

and administer the test requires about 40 to 45 minutes. Scoring is usually done by taking the total number of errors, but it can easily be scored, with the aid of a stencil, so that one secures the total errors for each variable, thus making four sub-tests.

With a group of music students this test correlated .671 (number 46) with a teacher's estimate. With the same group it had a self-reliability coefficient of .93.

Interval Discrimination. This test requires the comparison of two musical intervals played on the piano to determine if the last interval is longer or shorter than the first. There are 80 pairs of intervals to be compared. The first interval of each trial always begins in the region of the octave below middle C on the piano; the second interval is always given above the first interval so there is never any overlapping between the two. The intervals all proceed in an upward direction. Trials 40 to 80 are the reverse order of trials 1 to 40, with the intervals in each trial also reversed. This was done for the purpose of supplying a direct means for calculating the reliability by the test-retest method in the same sitting. Difficulty is increased in the first half of the test, whereas the last half becomes easier. The difficulty is based on the ratio of the difference in length of the two intervals. Thus in trial 1, C to B contains 11 half steps; its companion interval, F to A has 4 half-steps. Dividing 11 by four gives 2.75 which was taken as a measure of the difficulty, expressed as a ratio between the two intervals. Also the 40th trial contains 15 and 16 half-steps in the respective intervals; dividing 16 by 15 gives 1.06 which was taken as the measure of difficulty for this trial. Comparing these two trials, and others in like manner, it is found that the smaller the ratio the more difficult is the discrimination between them and the larger the ratio between two intervals the easier the discrimination becomes. It was found from the test results that the number of errors were distributed only

approximately as was anticipated from this scheme so there are exceptions to this arrangement.

A complete scheme of the test is given in Fig. 2. The test had a validity coefficient measured by teacher estimates of from .35 with age and training partialled out to a raw coefficient of .77. With musical subjects it was found to have a reliability coefficient, by correlating the first half with the second half (and correcting), of .83, while with a non-musical group it was .43.

Retentivity. This test was constructed to measure memory for the more elemental factors of melody than for melody itself as was the purpose in Musical Memory. It was developed primarily as a test of absolute pitch, or memory for isolated tones. In order to have a test of absolute pitch it is necessary to have an interval of time elapse between the first presentation of a pitch and the subsequent identification of it. The opportunity was taken to fill this lapse of time with memory for the length of intervals and with the memory for rates of time given by a metronome.

There are 20 trials each containing two sections. The first section is composed of three items which must be remembered to compare with the three corresponding items, arranged in the same order, in the second section. The items are:

(1) An interval played on the piano. The intervals are composed of every other trial in the Interval Discrimination test.

(2) A rate of time given by a metronome.

(3) Three notes played on the piano.

In the second section,—where there is an interval, it is to be compared (longer or shorter (L or S)) with the interval in the first section; where there is a rate of time, it is to be compared (faster or slower (F or S)) with the rate given in the first section; and where ONE note

FIGURE 2
SCHEME OF INTERVAL DISCRIMINATION

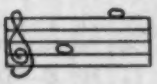
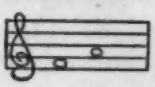
	a	b	c	d	e	f		a	b	c	d	e	f
1.	2.75	11	4	C– B	F– A	S	41.	1.06	16	15	D–#F′	A– C′	S
2.	2.50	5	2	G– C	A– B	S	42.	1.07	14	15	C– D′	G–#A′	L
3.	2.33	3	7	D– F	A– E	L	43.	1.08	12	13	F– F′	D–#D′	L
4.	2.25	9	4	E–#C	F– A	S	44.	1.09	12	11	C– C′	B–#A	S
5.	2.30	5	11	B– E	G–#F	L	45.	1.10	11	11	A–#G	F–#D	S
6.	2.00	3	6	C–#D	F– B	L	46.	1.11	10	9	E– D	C– A	S
7.	1.83	11	6	D–#C	E–#A	S	47.	1.12	8	9	D–#A	G– E	L
8.	1.80	9	5	F– D	G– A	S	48.	1.14	14	16	A– B′	D–#F′	L
9.	1.77	9	16	A–#F′	B–#D′	L	49.	1.15	15	13	E– G′	B– C′	S
10.	1.75	4	7	C– E	G– D	L	50.	1.17	14	12	C– D′	A– A′	S
11.	1.67	10	6	F–#G	G–#C	S	51.	1.18	13	11	G–#G′	E–#D	S
12.	1.60	16	10	D–#F′	A– G	S	52.	1.20	5	6	A–#D	B– F	L
13.	1.57	7	11	E– B	D–#C	L	53.	1.22	11	9	A–#G	F– D	S
14.	1.55	14	9	G– A′	B–#G	S	54.	1.23	16	13	E–#G′	C–#C	S
15.	1.50	4	6	C– E	G–#C	L	55.	1.25	4	5	G– B	A– D	L
16.	1.45	16	11	D–#F′	A–#G	S	56.	1.27	11	14	G–#F	D– E′	L
17.	1.44	13	9	E– F′	B–#G	S	57.	1.29	7	9	F– C	B–#G	L
18.	1.43	10	7	A– G	E– B	S	58.	1.30	13	10	A–#A′	F–#D	S
19.	1.40	5	7	C– F	A– E	L	59.	1.33	12	9	G– G′	E–#C	S
20.	1.37	11	8	F– E	B– G	S	60.	1.36	15	11	F–#G′	D–#C	S
21.	1.36	11	15	D–#C	F–#G′	L	61.	1.37	8	11	B– G	F– E	L
22.	1.33	9	12	E–#C	G– G′	L	62.	1.40	7	5	A– E	C– F	S
23.	1.30	10	13	F–#D	A–#A′	L	63.	1.43	7	10	E– B	A– G	L
24.	1.29	9	7	B–#G	F– C	S	64.	1.44	9	13	B–#G	E– F′	L
25.	1.27	14	11	D– E′	G–#F	S	65.	1.45	11	16	A–#G	D–#F′	L
26.	1.25	5	4	A– D	G– B	S	66.	1.50	6	4	G–#C	C– E	S
27.	1.23	13	16	C–#C	E–#G′	L	67.	1.55	9	14	B–#G	G– A′	L
28.	1.22	9	11	F– G	A–#G	L	68.	1.57	11	7	D–#C	E– B	S
29.	1.20	6	5	B– F	A– D	S	69.	1.60	10	16	A– G	D–#F′	L
30.	1.18	11	13	E–#D	G–#G′	L	70.	1.67	6	10	G–#C	F–#D	L
31.	1.17	12	14	A– A′	C– D′	L	71.	1.75	7	4	G– D	C– E	S
32.	1.15	13	12	B– C′	E– G′	L	72.	1.77	16	9	B–#D′	A–#F	S
33.	1.14	16	14	D–#F′	A– B′	S	73.	1.80	5	9	G– C	F– D	L
34.	1.12	9	8	G– E	D–#A	S	74.	1.83	6	11	E–#A	D–#C	L
35.	1.11	9	10	C– A	E– D	L	75.	2.00	6	3	F– B	C–#D	S
36.	1.10	10	11	F–#D	A–#G	L	76.	2.20	11	5	G–#F	B– E	S
37.	1.09	11	12	B–#A	C– C′	L	77.	2.25	4	9	F– A	E–#C	L
38.	1.08	13	12	D–#E′	F– F′	S	78.	2.33	7	3	A– E	D– F	S
39.	1.07	15	14	G–#A′	C– D′	S	79.	2.50	2	5	A– B	G– C	L
40.	1.06	15	16	A– C′	D–#F′	L	80.	2.75	4	11	F– A	C– B	L

a = ratio of the intervals.
b = number of half steps in first interval.
c = number of half steps in second interval.
d = musical notes of the first interval.
e = musical notes of the second interval.
f = correct answer; S = shorter; L = longer.
# sharp.
′ = there is a whole octave between the two notes.

is given, it is to be compared, or identified as none (0), first (1), second (2), or third (3) of the three notes given in the first section. The items for the trials are arranged in serial order, each appearing an equal number of times in the first, second, and third orders.

The metronome rate of 78 beats per minute was taken as the standard to be compared with the "rate of time" variables ranging from 164 to 88 beats per minute. The standard comes first in 10 trials and last in the other 10 trials, arranged in chance order. The metronome is allowed to beat either 4 or 6 seconds, always varying for the two sections of the same trial. (If the same number of seconds had been used it would have been possible merely to count the number of beats in each section and get the correct answer without paying any attention to the rate of the beats). One-half of the total number of sections containing metronome rates lasted for 4 seconds, the other half lasted for 6 seconds; this order was arranged by chance throughout the whole test, yet always varied for the two sections in the same trial.

To illustrate what the test requires with trial 1:—

there is first given a musical interval,  , on the piano; this is followed by the metronome beating at the rate of 164 beats per minute; and then three notes are given, . These are the three items in the first section, all of which must be kept in mind to compare with the respective items in section two, vs.:  is given and the testee must determine

if this interval is longer (L) or shorter (S) than the interval given in section one,—then a rate of time is given on the metronome of 78 beats per minute which must be compared to the rate given in the first section to determine if it is faster (F) or slower (S) than that in the first section; finally ONE note, is presented on the piano which is to be identified as one of the following: [1] none (0) of the three notes presented in the first section, [2] the first (1) of the three notes presented in the first section, [3] the second (2) of the three notes presented in the first section, or, [4] the third (3) of the three notes presented in the first section.

The subjects are given a record blank which appears like this for the first trial:—

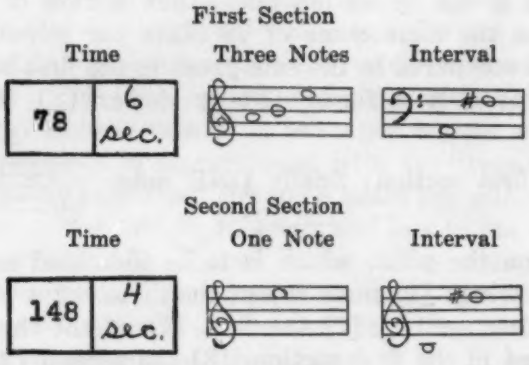
Interval	Time	Three Notes
L S	F S	0 1 2 3

They simply put a line through the correct letter, or figure on this form. In this example the correct way to answer would be:

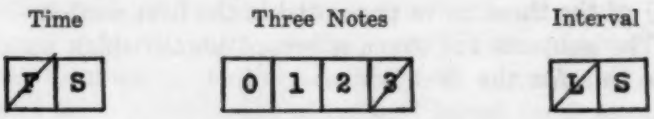
Interval	Time	Three Notes
L ~~S~~	F ~~S~~	0 1 ~~2~~ 3

The order of the items are the same in the two sections of each trial and the record blank has the order given for each, so no confusion results from having to remember the order of the items in addition to remembering the items themselves.

Trial number 5 is as follows:



The correct answer to trial 5 is:



The test as a whole yielded a validity coefficient of .466 with a musical group. The memory for tones (absolute pitch) item alone had an r of .394 with teacher estimate criteria. For the same group the reliability of the whole test was .76 and for absolute pitch item alone it was .61.

Intuition. This test attempts to measure some of the more subtle aspects of musical talent, through the medium of expression and feeling for musical form. It embodies the basic principles which have contributed in no small measure to the success of a well known academy of music in London where these principles have been carefully worked out and applied. The name comes from the belief that the musical response, or the desire to express feelings through music, is derived from an intuition which is in general opposed to intellect. The latter is, "merely concerned with the presentation of the idea."[1] It is therefore not a matter of training

[1] Trotter, T. H. Yorke. *Music and Mind*, p. 21.

which only helps to express this musical idea more adequately but an innate disposition which only requires correct development.

This musical response is best exhibited, and therefore measured, by the ability to supply an "answer" to a given unfinished musical theme. The successful accomplishment of this task is said to depend upon an intuitive feeling for three things:

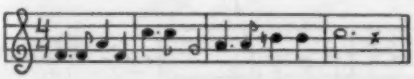
(1) phrase-balance, whereby the "answering" phrase should be of the same length as the initial phrase;

(2) key-centre, whereby the "answering" phrase should return to the central point (key) from which it started;

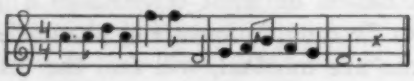
(3) time-balance, whereby the "answering" phrase should end in the same time as that in which the first began.

All good music does not conform to these three rules but they must be clearly felt or understood before successful deviations can be made from them.

The above three principles have been used as a test in practice for the selection of scholarship students, but it has never been standardized or written out so that all candidates were given the same material, and it was an individual test. A phrase such as



would be given and the candidate required to supply some answer. The quality of this answer is judged by these three principles and becomes fairly easy if one has had several years of experience in this work.

Thus  would be consid-

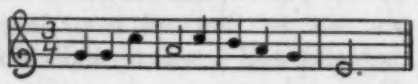
ered a good answer and

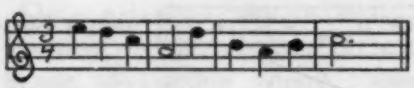
would be considered an inferior answer, for key-centre.

However the above procedure has many faults and the test of Intuition which is described here obviates them. Thus the test is made longer, the material is standardized, the scoring is made objective, and it is adapted for use with groups.

In order to insure that the test correctly embodied the three principles as already used in practice, the leading exponent of the school undertook to supply the necessary material. A test of each of the three "intuitions" was constructed so that both their individual and collective value, and interrelationships, could be determined.

Each of the three "intuitions" sub-tests contain 24 trials and by combining the three there are 72 trials for the entire test. Each trial consists of two given phrases without harmony. The testee is asked if the second phrase does or does not make a satisfactory answer to the first phrase. Thus, for the sub-test Phrase-balance

the first phrase might be 

followed by  . In order to

answer, one must compare the second phrase with the first to determine if it makes a "right" or a "wrong" answering phrase. In this case the answer is "right."

According to the theory upon which these principles are built, no explanation of what constitutes a good or a bad, or "right" or "wrong," answering phrase is necessary, for some, who are musical, will feel the difference

intuitively and those who are not musical will not discriminate the difference.

All the trials for each sub-test are given in immediate succession, the subjects being told which sub-test is to be given and examples of each are presented before the beginning of each sub-test.

The correlation between the teacher estimate and the total of this test, given in the school where all had had considerable training in these principles, was .304 (N—40) with age and training partialled out. With the same group each sub-test had the following correlation (age and training partialled out):

	r
Key-centre	.24
Time-balance	.29
Phrase-balance	.19

THE TACTUAL RECOGNITION OF EMBOSSED PICTURES BY BLIND CHILDREN

RALPH V. MERRY and FRIEDA KIEFER MERRY
Child Guidance Clinic
Dayton, Ohio

INTRODUCTION

The importance of pictures in the education of seeing children is well known, particularly in connection with the teaching of primary reading. Instructors of blind children are denied this valuable means of motivation and must discover other methods of making their subject matter interesting to their pupils.[1] This has been done to some extent through the use of natural objects and of models. Certain difficulties, however, are associated with the employment of these devices. Natural objects are so expensive that their extensive use is prohibited in the majority of schools for the blind, and cheap, bi-dimensional models appear to be of little or no practical value.[2] It can be asserted with truth, therefore, that no complete and satisfactory substitute for the pedagogical value of pictures exists at present in the education of the blind.

Furthermore, the adaptations of standard intelligence tests which have been developed for use with blind individuals are overweighted with verbal tests, since those of a non-language type frequently involve picture recognition.

[1] Merry, Frieda K., *Suggestions for Motivating Primary Braille Reading*, Series II, Curriculum and Instruction, No. 2, American Foundation for the blind, 1929.

[2] Merry, Frieda K., "A Study of the Merits of Animal Models Used in Teaching Blind Children", *The Teachers Forum*, II, January, 1930, pp. 12-13.

For some time teachers of blind children have been interested in the problem as to how much meaning the tactual recognition of embossed pictures has for their pupils. Publishers of braille books, also, have questioned the value of such pictures and have hesitated to reproduce many on account of the excessive cost. Although this subject has provoked much discussion, no experimental approaches have been made until recently.[3,4,5] Notwithstanding the fact that the first investigation[6] admittedly was of a preliminary nature, it indicated the importance of certain phases of the problem worthy of further research. Embossed pictures, made with an ordinary dressmaker's tracing wheel, of ten common objects (cup, chair, horn, table, fork, hat, shovel, wheel, house, apple) were presented to fifty blind children. Results showed that thirty per cent of the subjects made zero scores, and it was concluded, therefore, that the extent to which blind children can recognize embossed pictures tactually probably is very limited. No significant relationship was found between intelligence and the ability to recognize pictures (r .13 ± .09), or chronological age and picture recognition (r .28 ± .08). There was some evidence that children with partial vision were superior to the totally blind. The majority of the children who identified these pictures did so in sixty seconds or less. An exact analysis of the relative difficulty of each picture was impossible, but outstanding identifying marks, failure to comprehend perspective, and familiar-

[3] Merry, Ralph V., "To What Extent Can Blind Children Recognize Tactually Simple Embossed Pictures?" *The Teachers Forum*, III, September, 1930, pp. 2-5.

[4] Merry, Frieda K., "A Further Investigation to Determine the Value of Embossed Pictures for Blind Children", *The Teachers Forum*, IV, May, 1932, pp. 96-99.

[5] Merry, Frieda K., "An Experiment in Teaching Blind Children to Recognize Simple Embossed Pictures", *The Teachers Forum*, V, March, 1933, pp. 73-76, 78.

[6] Merry, Ralph V., op. cit.

ity or unfamiliarity with the objects depicted seemed to be pertinent factors in making the recognition easier or more difficult.

The first experiment herein described was undertaken to determine whether or not similar results could be obtained upon a larger number of subjects, and thus establish final conclusions upon a question which has been of so great interest to teachers of blind children. It differs somewhat from the preliminary study in that embossed figures (referred to as "designs") of an essentially bi-dimensional nature are included as well as pictures of tri-dimensional objects. In order to simplify conditions no subjects were included in this study whose vision permitted more than a perception of shadows.

The second experiment was carried out to determine how much the systematic teaching of perspective and the principles of graphic representation may improve the ability of blind pupils to recognize tactually embossed designs and pictures.

THE FIRST EXPERIMENT

Problems. The problems of the first investigation were:

1. How well do blind children recognize embossed designs and pictures by the sense of touch?

2. Are designs of essentially bi-dimensional objects easier to recognize than are pictures of tri-dimensional objects?

3. Do those who respond quickest make the highest scores?

4. What relationship do
   (a) general intelligence and
   (b) chronological age

bear to the recognition of embossed designs and pictures?

5. What is the relationship between age of blinding and recognition of embossed designs or pictures?

Subjects. Ninety-eight pupils, 49 boys and 49 girls, having no useful vision were selected from both the Upper and Lower Schools of Perkins Institution to participate in the experiment with *embossed designs.* They ranged in age from 6 to 24 years with a median of 14. While the lowest IQ was 61, the highest was 146 and the median was approximately 112. The age of blinding ranged from birth to 19 years with a median in the group where sight was lost during the first year of life.

Ninety-five pupils, 49 boys and 46 girls, having no useful sight were chosen for the experiment with *embossed pictures of a tri-dimensional nature.* Although the age range of these pupils was from 7 to 24 years, they were similar in intelligence and in age of blinding to the first group. All grades from the kindergarten through the high school were represented.

The procedure. Five embossed bi-dimensional designs were presented to the subjects in the following order (determined by chance):

1. Square
2. Cross
3. Triangle
4. Star
5. Circle

The size of these designs (5x5 inches), and the directions for administering the test corresponded to those used in the preliminary experiment. If the subject failed to respond within sixty seconds his score automatically became zero. If his answer was correct he received a score of two points for each design. One point was allowed for responses not absolutely correct but which showed an idea of the form, as for example, when the *triangle* was called a *tent,* the *circle* a *wheel,* or the *square* a *box.*

In the preliminary experiment it was thought that several of the pictures used probably were not well adapted to tactual recognition, hence other pictures were substituted. Their order of presentation, determined by chance, was as follows:

1. shovel
2. house
3. man's shoe
4. rabbit
5. chair
6. table
7. pitcher
8. stove-pipe hat
9. scissors
10. wheel

Two methods were used in administering these pictures, yielding two sets of scores. The subject was first asked to tell what the picture looked like,[7] and his response was recorded; then, if he failed to give a satisfactory reply, the examiner supplemented the directions by indicating the classification or use of the object represented in the picture. If this elicited a correct response it was scored separately. For instance, if the shovel was called a frying pan the examiner said: "Well, it might look like that, but suppose it is something you use in a garden. Now, what would you call it?" Further, if the table was called an elephant, the examiner suggested that although it might resemble an elephant, it really was a piece of furniture.

The time limit adhered to was one minute, the same as that used in the designs test. Two points were allowed arbitrarily for a perfect score, one for a partially satisfactory response (such as calling the *wheel* a *flower inside a circle*) and zero for an incorrect or incomplete answer.

Since results of the preliminary experiment had shown that the pictures were not all equally easy to recognize

[7] This evoked a much better response than the question, "What does that 'feel' like?" Apparently blind children use the term "look" in a perceptual sense, not in the strictly visual sense which it implies.

tactually, in this investigation the weight which each picture or designs score had in the composite rating was determined in accordance with the variability of the scores made on each picture or design. Thus, on the designs test the score for each subject comprised the total number of points for each of the five designs weighted according to its relative difficulty; *e. g.*, if a star was recognized correctly the two arbitrarily determined points were multiplied by three (making a total score of six points for this design), since it was three times harder to recognize than were the easiest, the circle and the square. In the same way two sets of composite scores were obtained for the embossed pictures. However, in the second method (where the examiner gave the subject a direct clue to the response) no partial credits were allowed, but the same system of weighting was followed.

In order to determine the better of these two methods the reliability of each was computed and was found to be .64 for the first method (wherein no suggestion was given) and .57 for the second method. Since it was impossible to administer the designs test at two different times and since the number of items (five) were so few, no attempt was made to figure its reliability.

Results of the first experiment. The highest attainable score on the *designs* test was 20, and was made by 33 (33.67 per cent) of the 98 subjects. The median was 17; there were no zero scores, and no sex differences.

Although the best possible score on the pictures test was 42 (when either the first or second method was used) the highest ranking obtained by the *first* method was 29, and 15 of the 95 subjects (15.78 per cent) received zero scores. The median was 5 and a statistically significant and reliable difference favored the boys, ninety-two chances in a hundred.

In using the *second* method there were no zero scores and one nineteen-year old girl (who lost her sight when

ten years of age) had a perfect performance. The median was 14 and the difference in favor of the girls was seventy-seven chances in a hundred.

The relative difficulty of the five designs was as follows: the square and the circle were easiest; the triangle came next in order of difficulty, and the cross and star were hardest to recognize.

Of the ten pictures the four easiest were: house, table, scissors, and wheel. The shovel ranked next in difficulty and the shoe, rabbit, chair, hat, and pitcher all were equally hard to recognize.

The relationship between the scores obtained on the designs test and the time required for the response was —.59 ±.04, while the correlation between scores and time on the pictures test was —.23 ±.06. The correlation between scores made on the designs test and IQ's was .50 ±.05 while that between IQ's and picture scores was .26 ±.06. The relationship of chronological age with scores on the designs tests was .35 ±.05 but with the picture scores was .12 ±.06. There was a correlation of .26 ±.06 between age of blinding and designs score and of .29 ±.06 between age of blinding and pictures score.

Conclusions from the first experiment. In answer to the problems undertaken in the first investigation it may be concluded that:

1. The ability of blind children to recognize tactually simple embossed designs and pictures is very limited, especially in regard to the latter.

2. Blind subjects are more successful in recognizing simple embossed designs or pictures of essentially bi-dimensional objects than they are in recognizing pictures of tri-dimensional objects wherein perspective is involved.

When the examiner furnished a clue to the recognition of the picture (as in the *second* method used) there was

a marked increase in score, especially with the girls. It seems, therefore, that if embossed pictures are included in braille books it would be desirable to supplement them with adequate description. The *first* method, however, is a more reliable test of a blind subject's native ability to recognize embossed pictures.

3. Generally speaking those who responded quickest made the best scores.

4. There is some relationship between general intelligence and ability to recognize bi-dimensional embossed designs, but this is negligible in the case of embossed pictures of tri-dimensional objects. The correlation between chronological age and ability to recognize either designs or pictures also is negligible.

5. In general the age at which vision was lost has practically no effect upon the subject's ability to recognize either embossed designs or embossed pictures.

The Second Experiment

The problem. The problem of the second investigation was to determine experimentally how much the systematic teaching of perspective and the principles of graphic representation may improve the ability of blind pupils to recognize tactually embossed designs and pictures.

Subjects. Eight younger children and seven older pupils were available for this experiment. Table I shows in detail the nature of these two groups. The younger children comprise Group I and the older children are classified as Group II.

On account of the small school enrolment *no useful vision* and *availability* were the chief criteria determining selection. Generally speaking, the younger group (I) was composed almost entirely of children of subnormal intelligence while the older group (II) was of average or better ability.

The procedure. Each child in Group I was instructed individually upon the forms of simple designs not exceeding a five-minute period three days a week from November 11, 1931 through December 16, 1931. Tests measuring the original ability of these children were

TABLE I

Data Concerning Pupils Participating in the Second Experiment

Initials of Subject	Sex	C.A.	M.A.	IQ	Grade
GROUP I					
M.R.	Girl	8-6	7-4	87	Kdgn.
I.C.	Girl	8-2	6-9	83	Kdgn.
D.S.	Boy	8-1	6-11	86	Kdgn.
B.K.	Boy	6-7	5-1	78	Kdgn.
N.B.	Boy	8-4	5-1	61	Kdgn.
A.G.	Boy	5-3	5-6	105	Kdgn.
A.F.	Boy	5-8	4-3	75	Kdgn.
A.S.	Boy	9-2	4-10	53	Kdgn.
GROUP II					
A.A.	Boy	10-0	9-4	94	3A
J.M.	Boy	11-2	12-10	115	4A
A.C.	Boy	11-9	14-0	119	4A
C.K.	Boy	10-7	11-3	107	3A
R.C.	Girl	11-6	13-4	116	5A
M.D.M.	Girl	14-0	12-10	92	6B
V.C.	Girl	12-6	13-0	104	4A

given during October, 1931, and final tests were made at the end of the period of instruction, followed by recall tests after one month's time had elapsed. From January 12, 1932 until February 15, 1932 each of these children was taught how to recognize embossed pictures of ten objects and two human figures. In a few cases the work was not completed until April on account of illness

and quarantine. A month after the close of the experiment the final two sets of the pictures were given as a delayed recall test.

The children in Group I were taught the designs according to their ease of recognition, *e. g.*, circle, square, triangle, cross, and star. A cylinder and an oblong were added. The method of teaching was as follows:

1. The examiner gave each child wooden models of the designs and identified them for the child if he failed to do so. *One* object only was presented in *one* lesson.

2. The child was shown the "picture" of the object and was told that that was how it looked to the eye.

3. Upholstery tacks in the form of one of the designs were inserted in a cushion[8] and the child was asked to make one like it. This procedure was followed for everything but the star which proved too difficult for them to construct.

4. At different sittings all the objects were placed before the child and he was asked to name them.

5. The next step consisted in letting the child hunt a certain design, as: "Find the cross."

6. By this time the child was expected to be able to match the wooden models with the embossed designs, as, for example, "Put the picture of the square with the cube."

7. The examiner frequently asked the child what design he had yesterday or the day before.

8. Occasionally, to add variety, the examiner said: "Show me the star," etc., instead of asking: "What is this?" (using a particular model).

9. Some of the perceptual forms given in Year IV, 2 of the Stanford Revision of the Binet-Simon Test[9]

[8] Commonly used in schools for the blind in connection with map work.

[9] Consult: Terman, Lewis M., *The Measurement of Intelligence*, Houghton Mifflin Co., Boston, 1916, pp. 152-154.

were embossed and used to determine if the child really knew the designs upon which he had been instructed or if he would be confused by suggestion. The correct forms were placed on the right, left, middle, etc., so that the child would not identify the design by its position. For example, embossed outlines of a square, oblong, and a square with a "bite" taken out of it ⊐ were placed before the child. The examiner said: "Give me the picture of the square. Be careful and don't let me fool you. Let's see if you're wide awake." Three trials, using different positions, were given. This technique was followed for each of the *original five* designs and the *two* supplementary figures.

10. By the fifteenth lesson each child was asked to identify the original designs. This was then followed by a recall test one month later.

From January 12, 1932 through February 17, 1932 the ten embossed pictures used in the first experiment were taught in their established order of difficulty, (from the easiest to the hardest), *e. g.*: *House, table, scissors, wheel, shovel, shoe, rabbit, chair, hat, pitcher.* Pictures of a *woman* and a *boy* were added to this list. In general the procedure was as follows:

1. *House*—The top and bottom of the house were indicated, and outstanding features were called to the subject's attention by such questions as: "What does smoke come out of?" "Where is the chimney?" "Put your finger on it." "Why do we have a roof?" "Show me the roof." "Where is the upstairs?" "Where is the downstairs?" "Find the windows." "Find the door." The examiner concluded by having the child point out all the features which were discussed.

2. *Table*—Child felt table upon which he was working. Examiner then explained to him how the top and

sides slanted in a picture. When a child responded to the question: "How many legs does a table have?" the examiner next showed him that they were all of the same size, but that this wasn't true in a picture. Drill was given to see if the subject understood *short* and *long*. In the picture he was told the front legs are long and the back legs look short, and was asked to find: "long legs;" "back legs;" "front legs;" "short legs," etc., thus associating the concept "front legs" with "long" and "back legs" with "short." Some children experienced so much difficulty in making the discrimination "long" and "short" that a lesson was introduced using the short and long Montessori steps.

3. *Scissors*—The picture and object were compared. "Show me the part that cuts." "Show me where you put your fingers." "Find the screw which holds the scissors together."

4. *Wheel*—The rim, hub, spokes, and spaces between the spokes were shown to the child using the wheels attached to a small wooden truck, before hunting these parts in the picture.

5. *Shovel*—Since no model of a shovel was available the children were told to examine the one used by the janitor. After discussing the shovel and its uses the examiner said: "Where is the handle?" "Show me the part that goes into the ground."

6. *Shoe*—The child identified the different parts of *his* shoe, after which the examiner asked him to show on the picture the toe, heel, laces, and the place where the shoe opens. The point was stressed that only one side of the shoe "showed" in the picture.

7. *Rabbit*—The children identified (on both a museum model and on a picture) the ears, tail, front paws, hind legs, eyes, and nose of the rabbit. Attention was called to his upright position in the picture and his reclining position in the model.

8. *Chair*—In teaching the concept of a *chair* a procedure similar to that used for the *table* was employed. The children were urged to look for different kinds of chairs and to notice the legs, backs, and seats.

9. *Stovepipe Hat*—Detailed discussion and description were necessary as no model was available. The crown, brim, ribbon, and stove-pipe effect were emphasized.

10. *Pitcher*—After identifying the spout, handle, bottom, top, and curved sides of a small china pitcher the children were asked to find them on the picture.

11. *Human Figures*—Parts of the body; front and side views; and peculiar positions (*e. g.*, as a boy with hands on his hips) were called to the children's attention.

When this series of lessons was completed each child was tested upon the original set of pictures, followed by a check list containing similar identifying elements. (The same material was given as a recall test a month after instruction had ceased in order to determine how much transfer in learning had occurred.) The check list comprised the following: a larger shovel; a different type of house (side instead of front view); a lady's slipper instead of a man's shoe; a rabbit in a horizontal rather than in an upright position; a chair with a cane seat; a table with a round instead of with an oblong top; a girl's hat as a substitute for the stovepipe hat; a large cup for the pitcher; a wheel of a different type; a fork for the scissors; and the pictures of a small girl and a man.

The technique followed for Group I was enriched considerably for Group II. Instruction was begun January 13, 1932 but was not completed in several cases until the middle of April. The lessons followed this outline in general:

1. Wooden models of a ball, square, cube, circle, cylinder, rectangle, oblong, equilateral triangle, pyramid, star, and cross were presented for identification.

2. Pictures of these models with and without perspective were used with the models themselves.

3. The children were required to make the cube, cylinder, and rectangle in perspective by sticking upholstering tacks in cushions.

4. The pictures were taught similarly to Group I supplemented by guided pencil drawings (child put tacks at certain places, as for example, at bottom of roof in drawing a house to guide him in making the sides, etc.). Outlines of many of the pictures were made with tacks upon the cushion over which a sheet of paper had been placed so that the child or examiner could trace the outline with a pencil. Free-hand folding and tearing of a shovel were used for one lesson. "Tack" pictures were made from both objects and embossed pictures as models, with and without perspective. In addition to the original embossed pictures the children:

1. Chose things they especially wanted to draw.
2. Drew a vase and a cup from models.
3. Drew an American flag from an embossed picture.
4. Studied a miniature of a house given to us by a Boston architect.
5. Made a man using plasticine.
6. Selected an object to be made from plasticine.

The tracing wheel was not used by the children to make pictures as it introduces too many complications, *e. g.*, making the picture backwards (on the reverse side of the sheet) so that it will appear normally on the front of the paper.

The procedure followed in giving check tests and recall tests was similar to that used with Group I.

Results of the second experiment. The percentage of improvement (or amount of transfer)[10] made by both

[10] The difference between the initial score (before teaching was attempted) and the score on the "recall" check list constituted the amount of improvement computed in terms of per cent.

groups of blind children in recognizing tactually embossed designs and pictures after intensive training is presented in Table II.

TABLE II

Amount of Improvement Resulting from Systematic Instruction in the Recognition of Embossed Designs and Pictures

GROUP I

Initials of Subject	Improvement (Per Cent) Designs	Improvement (Per Cent) Pictures
M.R.	25	—20
I.C.	70	40
D.S.	65	30
B.K.	55	20
N.B.	15	0
A.G.	95	—10
A.F.	55	0
A.S.	10	0

GROUP II

Initials of Subject	Improvement (Per Cent) Designs	Improvement (Per Cent) Pictures
A.A.	10	40
J.M.	0	10
A.C.	10	50
C.K.	30	40
R.C.	0	40
M.D.M.	10	20
V.C.	35	30

The results given in Table II show that instruction upon designs representing bi-dimensional objects yielded significant improvement for both groups, especially with the younger children. The older children made less im-

provement since, in general, they obtained higher initial scores.

There was less gain, however, in the ability of both groups to recognize embossed pictures of tri-dimensional objects, particularly with Group I.

Although exceptional blind individuals may be able to reproduce drawings employing the principle of perspective, it seems unwise to encourage any extensive teaching of pictures representing objects in three dimensions, unless used occasionally as a source of enrichment for mentally superior blind children. The question of how much real meaning such a procedure possesses for blind children in general must be considered carefully. Instead of giving reality to their ideas of things we may be increasing vagueness and verbalism, a criticism which any progressive teacher of blind children wishes to avoid.

Conclusions from the second experiment. Although the second investigation was carried on with but fifteen subjects, it seems unwise to expend any considerable amount of time teaching blind children how to recognize tactually pictures of three dimensional objects. If, however, the needs of the child demand instruction upon the identification of simple figures in two dimensions, time for such work seems justified.

General Conclusions

On the whole the findings of the preliminary investigation are confirmed by those of the first experiment and are borne out, also, by the results of intensive teaching in the second experiment.

It may be of value to make use of embossed designs of a bi-dimensional type in the education of blind children, but it is very doubtful if embossed pictures of tri-dimensional objects, wherein perspective is involved, possess any real meaning for children without sight even after systematic instruction.

THE MECHANICAL ABILITY OF SUBNORMAL BOYS

MARJORIE LOU PAGE*
Worcester State Hospital

Historical Background

Since it is customary to educate subnormal children along vocational lines the question has been raised as to the value of this type of training. Various studies have been made of the degree to which vocational training is assimilated by groups of institutional subjects. The present study concerned itself with the measurement of mechanical ability at a somewhat higher mental level, that reached by the "special class" system in the public school.

The present study is an outgrowth of the very comprehensive study of mechanical ability conducted by the Department of Psychology at the University of Minnesota under a grant from the Committee on Human Migrations of the National Research Council. (12).

Two fundamental assumptions have been in mind from the beginning of this study: 1. mechanical ability is

* Acknowledgments

The writer wishes to acknowledge invaluable aid from Dr. John E. Anderson, director of the Institute of Child Welfare of the University of Minnesota, who advised the study, and Professor Donald G. Paterson of the Department of Psychology who suggested the problem. The writer is deeply indebted to Dr. Mary M. Shirley for assistance in the preparation of this manuscript and for encouragement and assistance throughout the work. Acknowledgment is also made for the cooperation received from the other Research Assistants in the Institute of Child Welfare.

The writer desires to express her appreciation to Dr. Everly, Director of Research of the St. Paul Public Schools, for his unfailing encouragement and interest, and to the principals and teachers in the special class centers where testing was done.

measurable; 2. mechanical ability is unique with respect to other traits such as intelligence as measured. They state this theory thus (12—p. 21): "In brief, the theory holds that various degrees of success in all the important classes of human behavior correspond to compounds of relatively unitary traits, combined in various proportions, and that these unitary traits can be discovered and measured objectively.—From the point of view of the theory of unique traits the task of measuring mechanical ability must meet two theoretical demands. The first is that the tests be constructed to correspond with actual achievement in work generally regarded as requiring mechanical ability, identifying them as tests of mechanical ability only on the basis of correlations with outside criteria. The second demand is that measures be of mechanical ability and not of some other ability which may closely be related to it and which for that reason might possibly be confused with it. This demand can be met by constructing tests which, while giving high correlations with criteria of mechanical ability, will give low correlations with measures of other abilities and traits, for example, general intelligence."

In the present study an attempt has been made to test this assumption of uniqueness through the application of one of the tests developed to a group of one hundred subnormal boys. Aside from the theoretical importance of the study it is one of wide educational implications.

Since mental defect has been recognized many methods and theories of training have been developed. Most of these theories have recognized the inability of the subnormal to succeed under the prevalent modes of instruction and have attempted to find substitutes which would aid him in adjusting to life situations. In America the idea arose that since these children were unfit for academic training, physical training should be stressed. This theory and theories on the interrelation of mind and

body and transfer of training were soon combined. The result was an educational dictum, "train the body and improve the mind."

At the present time most large school systems follow the procedure of isolating children with intelligence quotients between fifty and eighty and teaching them in so-called "special classes." They attempt to fit the work to the level of mental development. In the large centers classification is usually on the basis of mental age and educational achievement. In addition to training in the fundamentals of reading, writing and arithmetic the industrial skills are stressed. The boys in the group tested received on the average one hour of this training a day. This amount may be compared with the hour-a-week schedule of normal boys of the same ages in the school system. Lightfoot (9) after a thorough analysis of a special class center voices the idea that manual training is stressed simply because they can not learn to do anything else. Baker (1) found the subnormal boys inferior in their manual training and thinks that they need the additional training for future adjustment in industry. It seems likely that the explanation of the present methods is to be found in a combination of the historic trends along with this demonstrated failure in the academic field.

However, the classes are taught in well equipped shops with special teachers and good materials. The practical question is whether these boys possess a sufficient amount of mechanical ability to profit by this type of training.

Many types of mechanical ability measures have been devised within the last few years. Most of them have satisfied to some degree the criterion of validity, namely high correlation with achievement in some task generally judged to require mechanical ability. Most of the work done in this field has been strictly utilitarian. The use of machinery and machine-made articles is rapidly increasing. It is estimated that at least forty per cent

of the gainfully employed in this country are engaged in work demanding mechanical ability. Investigators have attacked the problem with varying motives. Stenquist (13) was interested in the educational implications. Others have attacked it with a view to reducing wastes in industry. The Minnesota Study (12) was working toward an adult guidance program in immigration.

The tests devised may be classified into four main types: 1. mechanical performance tests; 2. tests of motor abilities involved; 3. verbal tests; 4. picture tests. (No attempt will be made to list all of the tests devised but the reader will find an adequate bibliography in the Minnesota Study (12).)

Several investigators have considered motor and mechanical ability of subnormal children as a training problem. Baker (1) found that skill and training in manual training were on a distinctly lower level for subnormal subjects. He stresses the importance of the training, however. Lightfoot (9) voices the idea that manual training is stressed simply because of other disabilities and does not report any remarkable result from the training. One current view is that it is not so much lack of ability as it is slowness of learning which makes the adjustment of the subnormal so difficult. Wilson (16) (17), Goddard (3) (4), and Link (10) have discussed this. Link (10) describes a training project in which subnormal girls were trained in factory work. He found that under careful supervision the girls could be taught simple processes. Stenquist (14) makes a similar plea for the salvaging of the subnormal.

Cowdery (2), working in the Whittier State School in California, made a comprehensive study of the relation between intelligence and success in trade training. He concludes that "general intelligence influences success in different fields to different degrees" but that "in no group is the degree or fact of success wholly dependent on intelligence." He proceeds to set up minimum mental

ages for success in each of the occupations taught. From these studies and from findings in industry we may conclude that there are places in industry which the subnormal worker can fill. It is quite possible that there are certain other positions in industry where his defect is an aid to his adjustment. The shop work taught in the school is hardly of this type. It is chiefly woodwork and drawing, both of which require skill beyond that required by the routine factory task. The question of the fitness of the subnormal for this type of training is a vital one.

In this connection it should be mentioned that time taken for learning is really of little importance providing that they are able ultimately to reach some degree of efficiency. Society can better afford long education than total support throughout life.

The test used in this study was the Minnesota Spatial Relations Test (12), a lengthened version of Link's (11). It is a complex formboard with a large number of forms. The original formboard was developed by Seguin in 1846. He used the board as an indication of mental development and as a training device. Most of these boards devised since then have been used as tests of intelligence and have been standardized by age norms. Jones (7) in 1913 recognized that the form board tested perception of form apart from general intelligence.

Link (11) was the first to adapt the test to industrial measurement. He included a complex board after an analysis of the work required of an assembler. Practically stated, he found that an assembler had to be able to "pick up a part, size it up, and then decide just where it belongs and how it shall be placed there." He called this ability the "perception of odd sizes and shapes." Scores on the test which he devised and foremen's ratings correlated to .72 for assemblers and to .81 for toolmakers and machinists.

The Minnesota Study (12) reported that the Spatial Relations Test was as successful as any used. It correlated .53 with a criterion of success in shop work and only .18 with intelligence. The question immediately arises as to why the test correlated so poorly with intelligence when the form board is a traditional intelligence test. There are two possible explanations, one in the nature of the test and one in the nature of the group tested. The first lies in the increased difficulty and length of the test. It is quite possible that these changes remove it from the mental test sphere. Furthermore, it is possible that a test which measures mental development at an early level may measure something different later in the developmental process. This possibility will be expanded more fully later.

METHODS

The present study is an attempt to measure the mechanical ability of a group of subnormal boys. The test used was the Minnesota Spatial Relations Test. One hundred subnormal boys from "special classes" in St. Paul were tested. This group will be compared with the one hundred seventh-grade boys tested by the Minnesota investigators (12) in their "experiment proper." This group will be referred to as the "Normal Group."

The mean chronological age for the normal group was 12.40 years; for the subnormal subjects 13.25. The subnormals ranged in age from 9 years to 15 years while the normal subjects ranged from 11 to 14 years. The standard deviation of the age distribution for the normal group was .80; for the subnormal group 1.45.

The normal group was given the Otis Self-Administering test of Mental Ability and intelligence quotients were computed. Each of the subnormal subjects had been tested with the Kuhlmann Revision of the Binet tests. The distribution of IQ's for the normal group showed a range of 50 points, from 75 to 125. The range for the

subnormals was from 50 to 85. The means were 101.11 and 73.1 respectively. The mental ages of the subnormal group ranged from 72 to 142 months (from 6 to 11.75 years). The median mental age was 114 months (9.5 years). There were no significant differences in IQ between the age groups. C.A. and IQ correlate .25; C.A. and M.A. .88 and IQ and M.A. .36. For the normal group these correlations were —.07, —.42, and .93. The differences noted between the groups are probably due to the greater variability of the subnormal group and and the homogeneity of the Seventh Grade Group. Other results of this sampling difference will be noted.

Grade location in the special classes is determined by achievement tests and mental age. The achievement tests are given frequently and classification is flexible. The subnormal boys were located in grades 1 to 5; all were retarded from two to four years compared with normal children of the same age. The mental and pedagogical retardation of these boys were about equal.

The occupations of the fathers of the subnormal subjects were classified according to the Minnesota Occupational Scale (6). Only 10 per cent were in the upper three classes, 18 per cent omitted the information and the remaining 62 per cent fell in classes V, VI, and VII. The occupations of the fathers of the normal subjects were rated according to the Barr Scale (15). The mean score was 9.54 with a standard deviation of 2.16. In terms of actual occupations the fathers of the subnormal subjects were chiefly laborers, semi-skilled, or unskilled while the fathers of the normal group were engaged in skilled types of work.

A scale similar to the Barr Scale was devised for grading occupations according to the judged amount of mechanical ability involved (12, p. 136). The mean score for the Junior High School group was 3.05±.79, for the subnormal group 2.92±.72. The difference between the means is the difference between the amounts

of mechanical ability needed by a small town station agent and a small dairy owner.

It was impossible to secure adequate data on the nationality of the subnormal subjects. The enrollment card had only the entry "Home Language;" if English was the sole language this was blank. If an additional language was spoken this was noted and we could assume that it represented the nationality of one or both of the parents. For the 59 per cent reporting English as the sole language there was no indication of nationality; 21 per cent were of North European extraction; 13 per cent South European; 3 per cent Negro; and 4 per cent Jewish. There was no reason to suspect that this group was not representative of the school population of St. Paul. There is a possibility that they differ somewhat from the normal group in this respect as the normal group was a Minneapolis school group. Lifson (8) found that Jewish boys were inferior in mechanical ability as tested by the Minnesota Spatial Relations Board. No other race studies have been reported in this field.

The differences and similarities between the two groups with respect to age, intelligence, and occupational status are summarized in Table I.

TABLE I

Summary of Description of Groups

	Subnormal Group	Normal Group
CA	13.25 S.D. 1.45	12.40 S.D. .80
IQ	73.1 S.D. 6.40	101.1 S.D. 10.80
Occupations of fathers	Unskilled labor	Skilled occupations
Mechanical ratings of father's occupations	2.92 S.D. .72	3.05 S.D. .79

The subjects were tested with Boards A and B of the Minnesota Spatial Relations Test (12). This test is a lengthened version of Link's (11) test. Each formboard

contains fifty-seven cut-outs. The reliability of the four boards given to the normal subjects was .84. Time considerations made it necessary to shorten the test in this study. Lifson (8) reported the odd-even reliability of Boards A and B as .87 for his subjects. The reliability for the subnormal group was .79. These coefficients are all corrected by the Spearman-Brown Prophecy Formula. In the main study (12) validity was figured by correlating scores obtained with the shop operations criterion. Since the reliability is very nearly as good as for four boards, and since in any case validity does not decrease as rapidly as does reliability it would seem that we might take the figure obtained in the main study, i. e., .52, as representative of the test given here. (Certain objections to this supposition will be discussed later.) The use of the two board form is justified by the main study (12, p. 216). A reliability of .80 and a validity of .52 justify the use of the test for purposes of group comparison.

All testing was done in rooms free from disturbance and interruption. Cases 1 through 50 were tested in May 1931; 50 through 100 in March 1932. The testing was done at a low table with the subject seated.

The directions were given exactly as in the main experiment (12, p. 311): "Board A is placed on the table, directly in front of the subject and flush with the edge of the table. The board is arranged so that the edge labeled A is the one farthest from the subject. The blocks are placed just beyond the empty board in the positions they have in Board B. This arrangement is most easily secured by placing Board B directly beyond Board A, i. e., from the standpoint of the subject, arranging the blocks in Board B, picking up Board B, and leaving the blocks on the table." The following directions were given to the subjects: "Put these blocks in their proper places on this board. They fit easily—do not force any of them. Use one hand. You may have to

turn them around but do not turn any of them over. Keep the black side of the blocks up. Go as fast as you can and get all of the blocks in Board. Ready, Go!" When the subject had placed the blocks in Board A this was pushed back, the blocks were removed as before, and Board B was presented.

Two scores were obtained for each board,—one in terms of the amount of time taken to place the blocks and one giving the number of errors made. The error score for each board is the number of times the subject tried to place a block in the wrong hole. Time scores are not only more objective measures but were found both in the original study (12) and by Lifson (8) to differentiate more closely within the group. The tests were interesting to the boys who regarded them as puzzles. Shyness was common but was overcome in every case before testing started.

The results were analyzed to find whether any differences appeared when the subnormal and normal groups were compared and to seek an explanation of any differences which might be found. In addition the relation of mechanical ability to such factors as age, intelligence, grade location and social status within the group was considered.

RESULTS

Gross Scores.—Scores on Boards A and B were combined to yield a total time score and a total error score. The comparison of the normal and subnormal groups in terms of gross scores is summarized in Table II.

When gross scores are considered the subnormal group is significantly inferior to the normal group both with respect to time taken for the task and errors made. The difference between the means are statistically significant in both cases. Only 25 per cent of the subnormal group equalled or exceeded the median time score of the normal group. In error scores only 7 per cent of the subnormal

subjects made as few errors as the median normal subject. The subnormal group showed a range of nearly five times as great as the normal in both time and error scores. This occurrence of extreme scores might explain the wide difference between the means. However, when medians are used the differences are diminished only

TABLE II

Comparison of the Groups in Terms of Gross Scores

	Subnormal Group		Normal Group
Mean Time	1194.5 seconds		72.5 seconds
Standard Deviation	758.4		194.4
D/o Difference		5.6	
Median Time	998.18		752.2
Overlapping	25%		
Mean Errors	61.4		16.7
Standard Deviation	48.4		7.3
D/o Difference		8.9	
Median Errors	45.4		15.9
Overlapping	7%		

slightly. The variability is not limited to a few cases but is a general wide dispersion.

It is apparent that the subnormal group is significantly poorer in this type of performance than the normal group with which it is being compared. Several factors may account for this inferiority. We shall consider the possible effects of the inequality of the groups in chronological age, in intelligence, in mental age, social status and in the amount of mechanical ability required by the father's occupations. The nature of the test itself will be considered as a possible factor.

Relation of Age to Success.—To what extent is chronological age an indication of ability to succeed on such a test? An analysis of the scores by age groups shows a consistent tendency in both groups for time scores to decrease with age. The tendency is not so consistent in the case of error scores. Boys below twelve years

of age made many more errors but the differences from age to age are not significant.

Chronological age and time scores for one hundred subnormal subjects correlated to .38±.06; chronological age and error scores to .46±.05. The Minnesota Study (12, p. 342) reports a correlation of .00 for time scores for the normal group. This disagreement of results is probably due to the greater variability of the subnormal group with respect to both factors. The age range for the subnormal group is very nearly twice that of the normals and the range in scores over four times as wide.

It is necessary then to seek further possible explanations of the inferiority in the other factors to be considered. It is possible that mental age and intelligence enter into the correlations given above. Age and mental age correlate to .88 and it is possible that the raw correlation given above represents rather the relation of mental age plus chronological age to test scores. Two techniques were used to investigate this possibility: the factors of mental age and intelligence quotient were held constant statistically by means of the partial correlation technique; and a method of limited sampling was employed to control them experimentally. When M.A. is held constant statistically C.A. and time correlate to —.36±.06. The partial correlation with IQ constant is .35.

It is usually more desirable to control such factors experimentally than statistically. For this purpose two groups of subjects with a narrow range in IQ were selected and the correlations computed for each group. Group A with 62 subjects with IQ's between 70 and 80 yielded a coefficient of .51; group B with 25 subjects and an IQ range from 60 to 70 yielded a coefficient of .36. We may conclude then that holding IQ constant does not materially affect the relationship between chronological age and time scores.

In a like manner 28 cases with mental ages of nine years were selected and the correlation was computed for this group. The coefficient was .002. Although the sub-groups were small and the variability unequal, these two methods of control suggest very definitely the rôle of the mental factor.

With respect to the relation of chronological age to time, we may conclude that what appears to be a clear and consistent relationship between the two factors tends to disappear when mental age is controlled either experimentally or statistically. The control of M.A. affects the magnitude of the relationship more than does the control of IQ. The actual influence of these mental factors will be discussed later. We do not find in the differences in chronological age between the two groups any explanation of the inferiority of the performance of the subnormal group.

Relation of Mental Age to Success.—Mental age and time scores correlated to .48, mental age and errors to .51 for the subnormal group. The relationship was somewhat less for the normal group ($r = .23$). Again the difference in the size of the coefficients may be due to differences in the variability of the two groups or to the nature of the test used. An analysis by age groups showed a consistent improvement in both time and errors with increased mental age.

When chronological age is held constant by the partial correlation technique the coefficient of correlation between mental age and time is .38; when IQ is held constant the coefficient is .45. Thus it is apparent that mental age, apart from chronological age and IQ is quite closely related to performance on the test for this group.

Relation Between Intelligence and Performance.—IQ and time scores correlate to .19 for the subnormal group and to .18 for the normal group. IQ and error scores

correlate to .22 for the subnormal subjects. When chronological age was held constant statistically IQ and time correlated to only .12, and with M.A. constant the coefficient was .02. We may conclude that IQ apart from mental age is not a factor of any importance in success on this test. The cases within the normal group with IQ's which fell within the range considered in the subnormal group all made scores above the mean for their group.

The Relation of the Social Factor to Performance.— Two ratings were made of the occupations of the father of each subject. The first rating was on a scale designed to classify occupations according to the amount of general intelligence required. The Barr (15) scale was used for the normal group and the Minnesota Scale (6) for the subnormal group. The second rating was made on a scale patterned after the Barr Scale and designed to rate occupations according to the amount of mechanical ability required.

For the normal group Barr scale ratings and scores on the test correlate only to .12. Because of the nature of the Minnesota Scale and the skewness of the group in this respect, correlations were not computed for the subnormal group. Analysis of the data in terms of medians for each occupational group did not reveal any relationship. It is unfortunate that a more detailed analysis of this factor can not be made. It would be most desirable to equate the groups more closely in this respect as well as in age. Two factors operate to make equality of the groups with respect to social status difficult. In the first place there is the positive correlation between intelligence and social status which has been reported by Goodenough (5) and others. This alone would account for the infrequent occurrence of children from the upper social classes in special class groups. In addition it is probable that where mental defect does

occur among the upper group, private schools and teachers are employed rather than the public special classes.

The mechanical ability rating of the father's occupation and the score made by the son correlated to —.08 for the normal group, and to .10 for the subnormal subjects. These ratings failed to correlate to any appreciable degree with the shop criterion in the main study. We may conclude that the father's occupation gives no indication of the score that the son will make on this test.

Discussion of Results

The data show a significant inferiority on the part of the subnormal subjects with respect to test performance whether time or error scores are considered. Several factors must be considered before one can evaluate the importance of this inferiority.

Probably the most important factor is the nature of the test used. The Spatial Relations Test is a complex and lengthened formboard. The formboard is a traditional intelligence test for children and subnormals. Attempts to devise adult performance tests of this type have been remarkably unsuccessful. It would seem from the result of formboard studies that they serve as intelligence tests up to a certain level, after which performance upon them is not an indication of mental level. There seems to be a very great possibility that the Spatial Relations Test may have been testing something different in the two groups compared in this study. This theory might be tested in two ways. A group of normal boys with mental ages equal to those of the subnormal group might be tested to discover whether the discovered relationship between mental age and performance is a function of the subnormality or the result of the mental level tested. A more satisfactory approach would include wider testing of the subnormal group with

a variety of mechanical tests for comparison with a normal group and the establishment of criteria of mechanical ability within the group for validation of tests.

It would be very desirable, if the sampling available permitted, to make this type of study employing a method of pairing in selecting the subjects. In this way the groups might be perfectly equated with respect to every factor except the ones under consideration, intelligence and mechanical ability. The limitations of this study made such a pairing impossible, and several sources of error in the resulting sampling should be mentioned.

In the first place the groups were not equated with respect to age. This greater age range in the subnormal group is reflected in all of the correlations in which age might conceivably be a factor; furthermore the inclusion of subjects younger and older than those in the normal group makes comparisons difficult. A wider sampling of dull children would make possible closer control of this factor.

Since there are apparently race differences in mechanical ability (8), more complete data as to racial extraction should be available for equating the groups.

The intelligence range studied offers great possibilities both from a theoretical and a practical viewpoint. They represent the group of subnormal individuals that we attempt to fit into the social scheme. The deviation from normal is not so marked as to make them unfit for education and employment. The nature of their suitable education and employment is an important question and much research is needed in this field.

Summary and Conclusions

One hundred subnormal boys enrolled in special classes in St. Paul were tested with the Minnesota Spatial Relations Test. The Minnesota Spatial Relations Test was developed by the Minnesota Mechanical Ability Project

(12) and was used in this study to obtain a measure of mechanical ability. The scores obtained by the subnormal subjects were compared with those made by the group of 100 normal seventh grade boys tested by the Minnesota investigators in their Experiment Proper. The results of this comparison are:

1. Subnormal subjects made significantly poorer scores on a test of mechanical ability than did normal boys.
2. Performance by the subnormal group was closely related to mental age and showed only a slight relationship to intelligence quotient.
3. In neither group is there a marked relationship between chronological age and scores when mental age is held constant.
4. Fathers' occupations give no indication of success of sons on this test.

Several factors were investigated for an explanation of the inferiority of the subnormal group. Of these the nature of the test seemed the most important. Further research, using other measures of mechanical ability, is necessary to determine the significance of these findings and to trace out in more detail the educational implications of the inferiority of the subnormal subjects.

BIBLIOGRAPHY

1. Baker, H. J.: Characteristic differences in bright and dull children. Bloomington, Ill.: Public School Publ. Co., 1927.
2. Cowdery, K. M.: Measures of general intelligence as indices of success in trade learning. Jour. Appl. Psychol., 1922, 6, 311-330.
3. Goddard, H. H.: Psychology of the normal and subnormal. Dodd Mead: New York. 1919.
4. ——: Human efficiency and levels of intelligence. Princeton University Press, 1920.
5. Goodenough, F. L.: The Kuhlmann-Binet tests for children of pre-school age. Minneapolis: Minnesota Press. Institute of Child Welfare Monograph Series. 1928.
6. Goodenough, F. L., and Anderson, J. E.: Experimental Child Study. New York: Century. 1931, 546p.

7. Jones, E.: A very good test of form perception and motor response. Psychol. Clinic, 1913, 9, p. 241.

8. Lifson, Nathan: A study of racial differences in mechanical ability. Unpublished study: University of Minnesota. 1931.

9. Lightfoot, Jessie: The study of a special class center. Psychol. Clin., 1928-29, 17, 190-199.

10. Link, H. C.: Education and industry. New York: Macmillan, 1923. 263p.

11. ——: Employment Psychology. New York: Macmillan, 1920, 435p.

12. Paterson, D. G., Elliott, R. M., Anderson, L. D., Toops, H. A., Heidbreder, E.: Minnesota Mechanical Ability tests. Minneapolis: University of Minnesota Press. 1930. 586p.

13. Stenquist, J. L.: Mechanical Aptitude tests. Yonkers: World Book, 1922.

14. ——: The case for the low IQ. Jour. of Educ. Res. Nov. 1921, 241-254.

15. Terman, L.: Genetic studies of genius. Vol. 1. Stanford University, Calif.: Stanford Univ. Press, 1925. 66-89.

16. Wilson, F. T.: Some achievements of the same mental ages but different intelligence quotients. J. Educ. Res., 1926, 14, 45-53.

17. ——: Learning of bright and dull children. Bureau of Pub., Teachers College, New York.

MAXIMUM AND MINIMUM INNER SPEECH IN READING

CHARLES BIRD AND F. S. BEERS
University of Minnesota

A survey of the literature has revealed no studies which evaluate the effect of maximum and minimum degrees of inner speech upon speed of reading. In regard to speed of performance, emphasis has been placed on comparisons between silent and oral reading. The results have yielded the well known fact that oral is definitely slower than silent reading, and the inference has been drawn that any degree of inner speech is likely to retard the reading process. The present study is an attempt to evaluate directly the differences in speed of silent reading under conditions which call for a maximum and for a minimum of inner speech.

To measure the effects of these two factors, it is necessary to have material which is simple and which therefore emphasizes speed; and, in addition, it is essential to have a check upon comprehension. The Chapman-Cook Speed of Reading Tests (Forms A and B) satisfy these conditions and furnish an excellent unit of measurement in terms of paragraphs, each of 30 words in length. Furthermore, as already demonstrated,[1] the two forms of the test are of equal difficulty provided Form B follows Form A immediately and provided the subjects have had no previous experience with these tests. It may be pointed out that unpublished data strongly suggest that the equality of the two forms is due to the warming-

[1] Paterson, D. G. and Tinker, M. A. *Studies of Typographical Factors Influencing speed of Reading.* IV. "Effect of Practice on Equivalence of Test Forms," *Jour. App. Psychol.*, 1930, 14, 211-217.

up effect resulting from practice in Form A. We have demonstrated in these data that when Form B follows Form A after an interval of two weeks, Form B is more difficult. For use with adults, it was necessary to reduce the time limit prescribed by the tests; in this study a time limit of one minute and three-quarters was used.

The tests were given to three different groups of adults, a total of 305 individuals. These three student groups were made up of classes in The American Institute of Banking, the Extension Division and the Summer Session of the University of Minnesota. For purposes of the experiment, the total number of subjects was divided into two major groups, the first one being made up of two sub-groups, to include members of the A. I. B. (Ia) and the Extension Division (Ib). These subjects, designated in Table I as Group I were required to read Form A with a maximum of inner speech and Form B with a minimum; Group II, composed of Summer Session Students, were required to read Form A with a minimum of inner speech and Form B with a maximum. This variation in order was used to compensate for whatever advantage might result from primacy of maximum or minimum articulation.

Since there are no means for determining mechanically the degree of inner speech exhibited by large groups of subjects, it was necessary to determine the degree of inner speech relatively and for this purpose to depend upon instructions and the extent to which the subjects could follow them. After Group I had completed the preliminary drill for the Chapman-Cook Test, Form A, they were given these additional directions:

"Read the paragraphs by pronouncing every word in inner speech. Pronounce each word by movements of the lips and tongue but do not read aloud. Read as rapidly as you can. Do no writing but merely cross out the one word which should not be in the second half of the paragraph. Remember both speed and accuracy count. Be sure, however, to read by pronouncing every word in inner speech, that is, by lip and tongue movements."

Immediately after Form A was completed, preliminary drill for Form B was given and then the following directions:

"This time you must read the paragraphs silently. Do not pronounce words by making movements of the lips and tongue. Read as rapidly as you can. Do no writing but merely cross out the one word which should not be in the second half of the paragraph. Remember both speed and accuracy count. Be sure, however, to read silently."

Group II were also given the tests in the same order, but minimum inner speech was required first. The results obtained are partially summarized in the table on page 185.

Inspection of Table I indicates that speed of reading is superior for all groups under conditions of minimum inner speech and that all of the differences, listed in the last column, are statistically significant. In the interpretation of these differences, it must be borne in mind that the formula for correlated measures has been used. The correlations between Forms A and B for each of the groups respectively were: Group I (A. I. B. and Extension students) +.76; Group Ia (A. I. B. students) +.83; Group Ib (Extension students) +.75; and Group II (Summer Session students) +.74. Although each group was comprised of men and women, it has not been deemed advisable to report in detail the data for each sex, but it may be observed in passing that the differences in favor of minimum speech hold consistently and significantly for each sex. Whether maximum inner speech is used for Form A or Form B, the average number of paragraphs read correctly is practically identical; and for all groups, it falls within the range of 14.28 to 14.88 paragraphs. The greatest discrepancy between minimum and maximum articulation occurs within Group II which read Form A under conditions of minimum inner speech. There is a suggestion in the results that members of this group were the superior readers

TABLE I

Differences in Speed of Reading Under Conditions of Maximum and of Minimum Inner Speech

Test Group	No. of Subjects	Average Age	σ Dist.	Aver. Formal Education in years.	σ Dist.	Mean, Form A (Paragraphs Read)	σ Dist.	Mean, Form B (Paragraphs Read)	σ Dist.	Difference Between Means	σ Diff.	$\frac{D}{\sigma \text{ Diff.}}$
I ---	152	27.8	7.25	12.28	2.07	14.8	4.24	16.38	3.99	2.08	0.23	8.96
Ia ---	40	27.25	6.26	11.18	1.44	14.38	3.51	16.8	4.02	2.42	0.36	6.83
Ib ---	112	28.0	7.56	12.68	2.12	14.28	4.47	16.22	3.96	1.94	0.28	6.83
II ---	153	25.94	6.99	14.53	1.41	18.33	4.48	14.88	3.45	3.45	0.21	14.12

of the two major groups; on the other hand, the order of minimum and maximum articulation may have had a selective effect which contributed to this superiority, yet the point is indeterminate from these data.

Earlier in this study, we suggested that performance depended in part upon the degree to which instructions were carried out. The consistency with which the differences favor minimum inner speech indicates that the adult subjects were able to adapt themselves significantly to the test situation. We do not contend, however, that all subjects met the test situation perfectly inasmuch as it ran counter to deep-seated habits. But apparently most subjects approximated the requirements since the superiority of minimum articulation is demonstrated by the average differences or by the percentage of each major group which showed greater speed. In Group I, 75.6 per cent of the subjects showed superiority during minimum articulation; in Group II 84.3 per cent. In contrast, 17.1 per cent of Group I showed greater speed when they read with maximum inner speech, and only 7.8 per cent of Group II.

In view of the growing interest in the effect of age upon performance in various academic situations, it might be well to mention the fact that the results from this study of reading indicate a slight superiority in favor of the younger members of both major groups. The Chapman-Cook test is primarily one of speed, and since recent investigations suggest that retardation is a characteristic of old people,[2] it need not be surprising to find that the test performance of the older members of our groups falls below that of the younger ones. If education is partialled out, the following correlations exist between age and speed of reading: Group I Form A —.16, Form B —.24; Group II, Form A —.38, Form B —.40. These correlations suggest merely a tendency for

[2] Sward, Keith. "A Study of Senescence, *Bulletin of Amer. Psychol. Association*, Thirty-Ninth Annual Meeting, 1931, pp. 6-7.

decreasing speed of reading to parallel increasing age; but, it should be kept in view that of the total 305 subjects in this experiment only 13.8 per cent had reached or exceeded the age of thirty-five years and that only two individuals had exceeded fifty years.

The analysis undertaken in this study is, of course, significant only for relatively easy reading material. Moreover, there is no guarantee that deep-seated habits of inner speech during the reading process can be materially reduced. Be that as it may, the test results emphasize the fact that most individuals who make an effort to keep their inner speech at a minimum level during the reading of simple material gain definitely in speed. This in itself is significant. Whether or not a gain in speed will result when more difficult materials are used for testing remains to be determined.

THE PREDICTIVE VALUE OF THE KUHLMANN-BINET SCALE FOR INMATES OF A STATE SCHOOL FOR THE FEEBLEMINDED

GRACE ARTHUR
Child Guidance Clinic, St. Paul, Minnesota

As psychological clinicians there are certain things we need to know about the scales we are using. We need to know the reliability of each scale for specific age ranges, for specific IQ ranges and for specific time intervals. We need to know the degree of reliability not only for the average case, but for the individual. In what situations are serious mistakes in diagnosis likely to occur? How can these be avoided?

Such questions as these will have to be answered for many different kinds of groups before any scale can be considered fully standardized. This report attempts to answer some of them for a single scale and for the population of one type of institution.

From the files of the Division of Research, Minnesota State Department of Public Institutions, retest data were obtained for 646 individuals, all but a few of whom were inmates of the Minnesota State School for the Feebleminded. The remaining few were children or siblings of inmates who were tested and retested by the institution examiners in order to determine whether they should have institutional or extra-institutional care. All the ratings used in this report were obtained with the Kuhlmann revision of the Binet-Simon Scale. Ratings obtained prior to 1912 were not included.

No data were included that were suspected of being affected by sensory or motor defect, extreme emotional disturbance or foreign language handicap. The number of retests for each individual range from one to six.

After all questionable data had been discarded, there remained 1207 retest ratings.

These were first divided into groups according to the time elapsing between test and retest. The formula used for finding the probable error is that of Otis.[1] In Table I the first column gives the time interval between test and retest. In the second column is the number of cases for each time interval. In the third column is the median difference between the IQ's obtained on the first test of the pair compared and those obtained on the second, and in the fourth is the probable error obtained from the median difference.

TABLE I

Interval	No. of Cases	Median Difference	P.E.
Less than 1 year	88	2	1.41
1 year	200	3	2.12
2 years	364	3	2.12
3 years	143	4	2.83
4 years	175	4	2.83
5 years or more	237	5	3.54

As one would expect, the median difference shows a slight but continuous increase with the increase of the time interval from less than one year to five years or more. Another change, but one not shown in the table, occurs with the increasing time interval. This is the increasing per cent of cases to vary in the direction indicated by Heinis' formula. For the group retested within a year as many cases earn a higher rating on the retest as decrease in rating. But as the time interval is extended a larger and larger proportion earn a lower rating on the retest than on the original.

The data were next classified according to mental age groups: 0 to 2.9, 3 to 6.9, 7 years and over seven. Those cases in which the first test yielded a mental age of 0 to 2.9 and the second mental age had fallen between 3 and 6.9 were placed in another group. Those for whom the

first mental age obtained fell between 3 and 6.9 while the second was seven years or more constituted still another. In Table II are given the values obtained from these groups.

TABLE II

Mental Age	No. of Cases	Median Difference	P.E.
0-2.9	93	4	2.83
0-2.9 to 3-6.9	34	8	5.66
3-6.9	444	3	2.12
3-6.9 to 7 and over	164	5	3.54
7 and over	472	3	2.12

According to these values ratings obtained within the age range 3-6.9 tend to run somewhat higher than do those falling within the range 0-2.9 or above seven. As the two transition groups together include less than one-sixth of the total number of retests considered it would be desirable to see whether the same results would be obtained for a larger number of cases at these levels.

Again re-classifying the data, this time according to IQ range, the following values were obtained. Each case was classified according to the IQ obtained on the first of the tests included in the comparison.

TABLE III

IQ Range	No. of Comparisons	Median Difference	P.E.
Less than 20	41	2	1.41
20-49	434	3	2.12
50-74	614	4	2.83
75-84	88	6	4.24
85-94	24	4	2.83
95-104	4		
105-115	2		

The IQ's above 75 are accounted for in part by the birth of normal children to feebleminded mothers who were pregnant when committed to the institution, in part by increase of ratings from definitely feebleminded to

borderline, in part by the commitment of individuals in the earlier days of intelligence testing on the basis of a diagnosis made according to the social definition of feeblemindedness, and in part by the testing of those closely related to inmates and suspected of abnormality. A single case of borderline intelligence that was tested five times supplies ten of these comparisons.

Throughout this report the commonly accepted classification according to IQ has been used: Idiot, 0-19; Imbecile, 20-49; Moron, 50-74; Borderline, 75-84; Dull, 85-94; Average, 95-104; Bright, 105-114.

From the three sets of averages presented in the above tables, it is evident that the Kuhlmann-Binet scale has yielded results of a high degree of reliability for this population. For the 1207 ratings the P. E. is 2.83.

But we need to know more than this. Do extreme variations ever occur? At what points does one need to watch for them?

The scores of each subject were inspected with these questions in mind. Each case was noted in which an increase or decrease occurred large enough to make a change in classification. In most of these the change was a drop from dull to borderline intelligence, or from borderline to feebleminded.

Three adult cases of this kind are counterbalanced by four others where mental age continued to increase after the point in chronological age had been reached beyond which the expected increase in intelligence is treated as negligible. In Kuhlmann-Binet ratings this point is 15.0 years. In each of these four cases the IQ had increased enough to raise the classification to the next higher group.

Conflicting tendencies were found also in a small number of cases of young children. Of two babies earning IQ's of 90 and 93 at the ages of nine and ten months respectively, one was definitely feebleminded at the age of nine years (IQ 34) and the other rated as borderline

(IQ 83) at the age of 18 months. Both of these changes occur in the direction indicated by Heinis' personal constant. On the other hand, six IQ's obtained by children with mental ages of less than three years improved enough upon retesting to change the classification: 1 from Imbecile to Moron; 1 from Moron to Borderline; 1 from Moron to Dull; 1 from Borderline to Dull; 1 from Borderline to Average (IQ 80 to IQ 102); and 1 from Moron to Bright (IQ 63 to IQ 111).

Inspection of these scores suggest the need of caution in diagnosing cases of delayed speech. It may be a symptom of feeblemindedness or it may constitute a true "retardation" that will be made up later.

To my surprise I found 6 examples of the "Goddard type" of case. These subjects reached a mental age on the first test that was not raised significantly on subsequent tests in spite of the lapse of years. An excellent example of this type of score is the following:

C.A.	M.A.	IQ
8-0	6-5	80
11-11	7-0	58
14-0	6⅞	49
15-11	7-0	47

Of this group, one dropped from IQ 100 to IQ 92, and another from IQ 110 to IQ 90. Whether subsequent tests will show these two to be examples of temporary retardation rather than permanent loss cannot be predicted. Both sets of mental ages fell within groups for which low P. E.'s had been obtained.

Of the 646 subjects whose test ratings are being considered, 590 obtained the same classification under repeated retests as that obtained on the first examination; 47 obtained classifications one step removed from that on the first examination; 7 obtained classifications two steps removed from that obtained on the first examination, and 3 obtained classifications more than two steps removed.

Of the last named, one was an adult retested within six months of the first examination. One was a case of delayed speech. One was a case with a rating of Dull (IQ 93) at 9 months and an Imbecile rating (IQ 34) at 9 years.

The changes as a rule occur downward rather than upward. The two situations most likely to present exceptions to this rule are found (1) in those individuals who continue to gain in mental age after they have reached a chronological age of 15, and (2) in young children who are handicapped by delayed speech.

From these results it would seem that there is very little danger of committing a normal individual as feebleminded upon the basis of a rating obtained with the Kuhlmann-Binet scale by a properly trained examiner. It is also evident that the prejudice in Dr. Kuhlmann's office and in the Child Guidance Clinic of St. Paul against recommending any individual for commitment as feebleminded upon the basis of results from a single psychological examination is of positive value in safeguarding the rights of the small percentage of patients whose scores may vary appreciably from one examination to the next. Moreover, in view of the changes in rating that occur with the development of speech it is doubtful whether any young child in whom speech is delayed should be committed as feebleminded until every effort has been made to encourage speech dvelopment.

SUMMARY

Twelve hundred seven retests of 646 individuals were obtained from the records of the Minnesota State School for the Feebleminded.

The P. E. obtained by the use of the Otis formula was found to be 2.83.

Five hundred ninety cases obtained the same classification on repeated retests as that obtained on the first examination.

Forty-seven cases obtained classifications on retests that were one step removed from that obtained on the first examination. Of these 25 moved up one step, 21 moved down one step.

Seven cases obtained classifications on retests two steps removed from that obtained on the first examination. Of these, five earned lower ratings on the retest. The two who moved up two steps in classification were infants.

Three cases obtained classifications on retests that were more than two steps removed from that obtained on the first examination. Two of these were infants. One was an adult retested within six months of the first examination.

From these data it appears that the Kuhlmann-Binet Scale in the hands of properly trained examiners furnishes results of a high degree of reliability in diagnosing the feebleminded. For the sake of the occasional case it is safer to share the prejudice of the Minnesota State Research Bureau and the Amherst H. Wilder Child Guidance Clinic of St. Paul against making a recommendation for commitment to state guardianship as feebleminded on the basis of a single examination. This is especially necessary to avoid mistakes in the diagnosis of young children in whom speech is delayed.

FACTORS INFLUENCING THE AMOUNT OF COPY READ IN MAGAZINE ADVERTISEMENTS[1]

STANLEY G. DULSKY
University of Chicago

INTRODUCTION

In spite of the large amount of research done on psychological problems in advertising one very important line of investigation has been totally neglected. Most of the experiments performed in psychological laboratories have been concerned with two major topics, namely, attention and memory. But a very important practical problem is: How much copy is read in an ad and what are the possible factors influencing the amount read?

The question has never been answered experimentally. Men are paid to write copy, and they do, prolifically. How many people will see and actually read this copy the advertising man does not know; he tacitly assumes that it will be read by a goodly number of people.

The lack of a practical technique seems to me to be the primary reason for the neglect of the above-mentioned problem. As early as the "nineties" psychologists had devised methods of measuring memory and attention. These methods were then applied to practical problems with the result that there is a mass of information on the attention and memory value of advertisements. The writer believes that with the method used in this experiment he has added a new tool to psychological research in advertising.

[1] The writer wishes to express his appreciation for the kind assistance rendered by Dr. F. A. Kingsbury.

ORIENTATION

The one scrap of information we have concerning this problem is the result of a questionnaire by Starch.[1] In a personal interview he asked 603 persons in Boston and immediate environment a number of questions concerning their habits in advertisement reading. He found that the average number of ads in a given issue of a magazine read entirely through, per person, was three. Attention value of ad, layout, unusualness, size, etc., interest in product advertised, and attention and interest value of illustration were the main reasons for reading the copy.

The results of the above questionnaire do not furnish us with reliable information. The data are open to the objections raised against questionnaires in general. The present investigation was undertaken, therefore, to determine experimentally how much of the copy is read and the factors that influence the amount read.

PROCEDURE

The subject was seated at a table with the experimenter. A clean copy of a magazine, November 8, 1930, issue of the *Saturday Evening Post*, and a large, soft blue pencil were placed before him. The following typewritten instructions were handed to the subject.

> This is an experiment to see how different people read advertisements. It is *not* a test of your ability or memory. When the signal is given, open this magazine to page 52, and beginning there, look through the following pages in your *customary manner*, just as you would in looking over a magazine in your own home. As you read the advertisements make a vertical mark with this pencil along the right margin of the line read. *Read as much or as little as you care to.* If you only read the headline put a mark on the right margin of it and go on to the next ad. If you read the print below and stop in the middle of a sentence put a mark to the right of the last word read. There

[1] Starch, Daniel—*Principles of Advertising*. 1927. A. W. Shaw Co., Chicago and New York.

is no time limit. Make yourself comfortable and at ease before starting to look over the magazine.

When the subject had finished reading the instructions and the experimenter was satisfied that they were understood the subject was told to open the magazine to page 52 and begin reading. After the subject had looked through the magazine to page 86 he was told to turn to page 124 and continue from there. When the subject had read page 144 he was told again to skip to page 164. He then read through page 166 and was told to stop.

The experimenter then presented the subject with the following instructions:

Check on the line shown your estimate of the *amount* you read as compared with the amount you usually read.

many times less	much less	somewhat less	same	somewhat more	much more	many times more

Check on the line shown your estimate of the *time* you spent in reading as compared with the time you usually spend in reading.

many times less	much less	somewhat less	same	somewhat more	much more	many times more

The experimenter sat at the opposite end of the table with a small partition of wood between him and the subject. Behind this partition, and unknown to the subject, the experimenter recorded the time spent on each ad. The second hand of a pocket watch was used. Finer differences in time were not deemed necessary.

Seventy subjects served in this experiment. The group included forty-two college students, sixteen high-school students, eight business men, and four housewives.

RESULTS

1. Method

Obviously, the value of our results depends on the efficacy of the method. Even though all laboratory situations are artificial we believe that the marking method reproduces a natural atmosphere and attitude as far as any experimental situation can. Many people when they read hold a pencil in hand and follow the printed matter. This situation was duplicated in our procedure. Only two of the seventy subjects reported that they were distracted.

To ascertain how efficacious the marking method was the following plan was adopted. Numerical values were assigned to the statements which the subjects checked at the close of the experiment. "Somewhat more" was given a value of 1; "Much more," 2; and "Many times more" was given a value of 3. The statements, "Somewhat less," "Much less," and "Many times less" were given corresponding negative values. "Same" was given a value of zero. A check-mark on the line between two statements was given a value lying midway between the two numerical values of those statements. For example, if a mark were placed on the line between statements "Somewhat more" and "Much more" it was given a value of 1.5. In this manner the positive and negative values were computed for time and amount for all the subjects.

TABLE I

Estimates of Amount Read and Time Spent

	Amount			Time	
Value	No. Ss.	Amt.	Value	No. Ss.	Amt.
Plus	25	30	Plus	23	27
Minus	20	20	Minus	28	29
Zero	25	0	Zero	19	0

The above table shows that 25 subjects reported that they read more than they usually do, 20 reported they

read less, and 25 reported that they read as much in the experiment as they usually do. The 25 who read more read on the average "Somewhat more" (computed by dividing 30 by 25 which gives 1.2, or a little over 1, the numerical value of the statement "Somewhat more"). Similarly, the 20 who read less read on the average "Somewhat less." The average deviations are small. Therefore, the tendencies of 25 individuals to read a little more are balanced by the tendencies of 20 individuals to read a little less; so our seventy subjects, as a group, reported that they read approximately as much copy in this experiment as they normally do in an everyday situation.

The time values are worked out in a corresponding manner: 23 spent more time, 28 less time, and 19 the same. The people who spent more time averaged "Somewhat more" and the people who spent less time averaged "Somewhat less." Again the two tendencies balance.

We conclude, therefore, that the averages obtained by the marking method are sufficiently reliable (though individual averages may not be) so that they can be taken as approximations of the amount of copy people read and the time they spend in looking at ads.

2. *Amount Read*

The data used were obtained from 40 ads distributed as follows: 11 full-page no color, 10 full-page color, 9 2-column, and 10 1-column ads. The first five ads, pages 52-56 inclusive, were not included in the scoring. They served as "shock absorbers," *i. e.*, exercises to allow the subject to become familiar with the procedure.

First, let us define the word, copy, as we shall use it in this report. Copy as generally defined includes all the printed matter in an advertisement. We are using the word in a more restricted sense, to designate the text,—or all the wording exclusive of headlines, sub-headlines, price, product and brand name, slogans, trade-marks, and

fine print contiguous to and explaining or describing illustrations.

Table II gives the results obtained from the four groups of ads. Column 1 gives the type of ad, *e. g.*, full-page no color, full-page color, etc.; columns 2, headed "Av. No. Readers," requires an explanation. Not all the subjects read the copy. Some merely read the headline and passed on. Only those who read part of the text are classed as readers. The headings of the other columns are self-explanatory.

TABLE II

Class of Ad	Av. No. Readers	No. Wds. Read per Reader	Av. No. Wds. in Text	% of Text Read by Readers	% Read by All
FPNC	12	96.02	270.27	39.06	5.9
FPC	12.2	85.5	194.5	50.45	10.15
2-col.	10.33	86.37	206.11	44.28	6.5
1-col.	8.3	48.38	139.9	36.87	5.32

We have now an answer to the question: How much copy is read? Taking all the results together we are justified in saying that about 15 per cent of all the people who see an ad will read some of the text and that this 15 per cent will read an average of 40 per cent of the text. Or, we may say that from 5 to 10 per cent of the copy will be read by everyone who sees the ad, under conditions comparable with those in our experiment, *i. e.*, when attention is called to the ads and subjects are expected to look them over.

The slight superiority of full-page color ads as shown in the column, "% of Text Read by Readers" is not reliable because only two ads caused this superiority. No conclusions may be drawn concerning the influence of size of ad on amount read because the deviations and consequently the probable errors of all the averages are so great. With more ads and better controls significant differences may appear.

3. Time Results

Table III gives the time (in seconds) spent on each group of ads for readers and non-readers and also the total average time.

TABLE III

Class of Ad	Av. Time Spent by Readers	Av. Time Spent by Non-Readers	Total Average Time
FPNC	25.9	7.33	10.88
FPC	24.46	8.41	11.67
2-col.	22.3	6.65	9.56
1-col.	15.1	4.97	7.01

A rank-order correlation was computed between percentage read and time spent. It amounted to .21. The average time spent by non-readers was subtracted from the time spent by readers. This difference in time should be, theoretically, the time spent by the readers in reading copy. We then correlated this time difference with percentage read and obtained almost the same coefficient as before, .20. It seems that though a subject spends much time on a particular ad, this time is not utilized entirely for reading. The subject might read a line, then refer back to the picture or headline and then continue. It indicates clearly that time cannot be safely taken as an index of amount read.

Factors Influencing the Amount of Copy Read

1. The Age Factor

The age of the subject was found to be a very important factor influencing the amount read. Out of our group of 70 subjects, 15 were under eighteen years of age. While our young group, those under eighteen, formed 21.4 per cent of the total number of subjects, they read only 12 per cent of the total. Sixty-seven per

cent of the adult group read some copy; only 29 per cent of the young group read any. We may conclude, then, that fewer adolescents read copy than adults, and that when they do read they read much less than adults. This is possibly a reflection of the more fully established buying habits and interests of adults. Table IV gives the results for each group on each set of ads. The differences are consistent and marked.

TABLE IV

Amount Read by Young and Adult Groups

	Young Group		Adult Group	
	Per Cent Readers	Per Cent Read	Per Cent Readers	Per Cent Read
FPNC	7	5	19	95
FPC	10	19	18	81
2-col.	5	10	17	90
1-col.	7	15	13	85

2. *Right vs. Left Pages*

Since Nixon has asserted that left-hand pages have superior attention value we have examined the data to see if position has any effect on the amount of copy read. Our data included 21 advertisements on right-hand pages and 19 on left-hand pages. Table V gives the results; the size of the ads, the means of the percentages read for left and right-hand pages, the difference between the means, and the probable error of the difference.

TABLE V

Size	R. Page Mean	L. Page Mean	Difference of Means	P.E. of Diff.
FPNC	34.79	39.44	4.65	4.88
FPC	45.42	57.99	12.57	6.24
2-col.	40.07	46.39	6.32	4.8
1-col.	31.58	40.39	8.81	7.07

For each group of ads the left-hand pages show a larger mean than the right-hand ones. Although the

probable errors of the differences are too low for statistical reliability, in all four classes there is a superiority shown for left-hand pages. With more ads in each group the differences would probably be more striking and the reliability greater.

3. Number of Words in Text

A rank-order correlation was computed between the number of words in the text of the ad and the percentage read. The coefficient obtained was —.32. There is thus a slight tendency for a smaller percentage of copy to be read in the more wordy ads. A rank-order correlation coefficient of .62 was obtained between the absolute number of words read and number of words in the text.

4. Amount of Information

The writer was of the opinion that the more information given by the headline and picture about the product the less would be read. Therefore, ten judges were asked to rate the 40 ads as to the amount of information given to them about the product by merely looking at the headline and picture. Correlations were computed between amount of information and percentage read. In no cases were the coefficients above .20.

5. Sex Differences

The group of 70 subjects consisted of 28 women and 42 men. Our numbers were large enough to study sex differences if any existed. The percentage of each ad read by the women was compared with that of the men. There were no marked differences. We may conclude that sex is not a factor influencing the amount of copy read.

6. Size of Print

An attempt was made to see if the size of the print in the copy had any effect on the amount read. Since most

of the print was 10 and 12 point type, the differences were not sufficiently marked to be of significance. In a group of ads showing wider typographical diversity, differences might appear.

SUMMARY

Seventy subjects were asked to read the advertisements in a magazine in their usual manner and to mark whatever they read (marking method). The results may be summarized as follows:

1. The marking method is an efficacious technique in studying certain advertising problems.
2. The average person reads 5 to 10 per cent of the copy of an ad.
3. The average person spends about 10 seconds looking at an ad.
4. There is no correlation between amount of copy read and time spent in looking at an ad.
5. Age was found to be an important factor influencing the amount read. Young subjects, 14-18, read far less than adults.
6. There is a marked tendency for subjects to read more copy on the left-hand pages.
7. There is a correlation of —.32 between the number of words in the text of an ad and the percentage read. There is a tendency for a smaller percentage of copy to be read in the more wordy ads.
8. Size of type, amount of information imparted by headline and pictures concerning the product, and sex, did not, under our conditions, affect the amount read.

THE INTERVIEW AS A RESEARCH AND TEACHING DEVICE

GEORGE W. HARTMANN
Pennsylvania State College

Progress in psychological research as in all other fields of scientific activity is to a large extent dependent upon the discovery and application of new tools of investigation. One has only to reflect upon the incomparable service which the microscope has rendered the biological disciplines to appreciate the significance of this statement. Imagine the present status of the social sciences had not Galton and Pearson provided them with the enormously useful coefficient of correlation! Certainly the greater bulk of contemporary educational research would be absent had not the testing movement emerged about the turn of the century.

The specific purpose of this paper is to call attention to the advantages of a simple and readily available instrument of research commonly known as the interview. The neglected possibilities of this modest device make one wonder why it has not been used more extensively to broaden the positive accumulations of educational theory. Perhaps it lacked the halo of apparent precision so much favored by the early enthusiasts in mental measurement; perhaps it was thrust aside because it seemed too clumsy by comparison with the much more convenient printed questionnaire; but most likely it was simply not considered because no one thought that anything intellectually valuable could be obtained by such an ordinary procedure. A similar situation has recently given rise to Gestalt psychology, largely because the older experimentalists uncritically felt that nothing problematic was presented by the existence of perceptual configurations.

To be sure, it is now more than a decade since Charters demonstrated the serviceability of this technique in a brilliant and original paper on "The Collecting of Unrecorded Specifics"; (2) but with few exceptions, its development as a standard piece of methodology in the various divisions of educational interest has been left in the hands of Charters and his more intimate associates.

Like most of its investigational tools, the interview is not a device peculiar to education alone. As an immediate and practical fact-finding agency, usually devoid of any scientific purpose, it is widely used in law, business, and journalism. With this *casual* as opposed to the *systematic* employment of the interview we are not for the moment concerned. More pertinent illustrations of its effectiveness can be gathered from the related spheres of psychology—particularly the clinical variety,—economics, anthropology, and sociology. Workers in the last two sciences have especially noted precautions to be observed in making field investigations lest they be misled by the acquiescence of those whose customs and beliefs they are studying. Research workers in education and related practical fields would profit by a greater familiarity with the more extensive experience of these specialists with the benefits and limitations of the interview as an instrument for collecting data.

I do not wish to give the impression that the use of the interview for educational studies is a wholly novel affair or that it has not as yet been adopted with marked success. Much of the outstanding work in job analysis and curriculum construction is founded upon a skillful use of the interview, and a number of recent handbooks for training graduate students in the ways of research recognize it as a method coordinate with parallel group experimentation, documentary analysis, etc. What I am pleading for is an *extension* of the method to a wider variety of problems, particularly to those problems which so far have not been attacked by *any* technique or to

those issues which promise to yield more fruitful results through the interview than through the use of other procedures. Lest some persons should feel that this proposal involves the substitution of a cruder tool for more highly refined ones it is only necessary to remind them that the individual mental test is merely a highly standardized interview; and it is doubtful if education or personnel management can boast any more accurate instrument. The decennial census of the United States, itself one of the finest sources of statistical information available, is built primarily upon data collected through interviews.

To what special problems of the educator could the interview be applied with some prospect of success? One does not need to glance far over the field before being persuaded that it is well-nigh universally applicable. The reason is simple: most educational work involves a question of human relations. The interview—or conversation with a purpose, according to the apt description of Bingham and Moore (1)—arises out of as many emergencies as school life itself presents. Whether he realizes it or not every administrator or teacher meets these emergencies in at least four general forms: 1, in vocational and educational guidance; 2, in placement; 3, in student discipline; and 4, in personality adjustment. Every schoolman with a minimum of experience realizes that each of these situations fairly bristles with controversial material.

In many investigations it is undoubtedly true that the interview is more valuable as a supplementary device than as a primary instrument, but even where it legitimately occupies this subordinate position it can generally be relied upon to insure a more intimate and personal acquaintance with the raw data. Often in these cases the judiciously guided interview permits a deeper insight into complex mental mechanisms than could be obtained through any other procedure. For example, the inci-

dence of truancy and cheating at different class levels might roughly be predicted by a regression equation containing variables such as severity of the teacher, intelligence score, occupation of father, etc. Nevertheless, an understanding of why these relationships are influential can often be gained only through direct and sympathetic inquiry. In such an important problem as the factors affecting teachers' morale, the interview is the only reliable technique available for obtaining this information. Industrial psychology has already two classic studies of this sort,—one by Houser (8) on employers' attitudes and another by Hersey (7) on the cyclical changes in the feeling tone of industrial workers. Other sample investigations which could properly be approached along the lines of the interview are: the determination of the qualities of a good department head; the most effective procedures in reprimanding subordinates; the methods used by successful teachers in making students work up to capacity; and of course, all the numerous phenomena of personality which are referred to the psycho-educational clinic for correction.

Despite the fact that the questionnaire has fallen into disrepute during the last decade, many persons will ask, Why isn't it just as good as, if not superior to, the interview? In his "Curriculum Construction" (3) (p. 135), Charters maintains that in *all* cases the oral questionnaire is preferable to the written form. I should say in *most* cases this is true, since Davis's study of women's sex life found the questionnnaire better for her purposes —but this was probably due to the peculiar strength of the social disapproval involved in her investigation. The chief objection to the interview is the labor required of the research worker himself and the element of suggestion involved in the situation (although this last is by no means absent from the printed form). The advantages, however, are many. Misconceptions as to meaning can be cleared up; supplementary questions will elicit

more definite answers; the person interviewed will supply the material more readily because the mechanical effort of writing demands too much of his time; and last but not least, from the standpoint of adequate sampling, more questions are answered and the separate replies tend to be more detailed. Some have maintained that a quantitative interpretation is impossible since the response cannot be scored; but there is no reason why much of the data thus secured cannot be converted into tabular form before further treatment is undertaken. Moreover, the mere existence of Snedden's (11) clever disguised intelligence test and O'Rourke's (10) method of measuring resourcefulness is sufficient refutation of this charge.

The careful research worker will always wish to know the validity and reliability of any research instrument and it must be acknowledged that the interview does not fare so badly when tested by these criteria. The correlations found by Clark (4) between interviewer's estimates of students' ability and their actual marks ranged from +.66 to +.73 and certainly indicate that interviewing possesses some value. As to reliability, Moore found in his study of textile workers on strike a correlation high in the .90's between the reports obtained from the first sample of 80 strikers and a larger group of 500. Of course, the reliability of different types of individual questions is another matter. Here common sense dictates that one should not ask for material which can better be obtained elsewhere from documents. A sense of the fitness of things will also determine whether a written record of the interviewee's responses should be made *during* or *after* the conversation. If the material is free from emotional loading—as in the case of data pertaining to practices in extra-curricular activities, methods used by competent scientists in developing "hunches," etc.—and if the person interviewed is in sympathy with the purpose and manner of the investigation,

then no harm is done and much is gained by immediate recording of the responses; but where intimate and delicate inquiries are involved, authorities suggest postponing the write-up until the close of the session. The memory flaws possible in this instance are less serious than the injury to spontaneity of answer which would otherwise occur. The interview thus duly guarded has many uses, particularly in the field of diagnosis and treatment, but from the research standpoint it is chiefly valuable for ascertaining facts of personal experience, opinions, and attitudes which only the individual himself can supply. However, since these attitudes may constitute, in the judgment of many sociologists and philosophers, the most important outcomes of the educational process, the urgent necessity for examining them through the most appropriate channel is obvious.

There is another possible function of the interview, bearing less of a research than a pedagogical character, which should not be overlooked in any inquiry because it is found in connection with the all-important problem of improving teaching efficiency. Some exploratory investigations at the Ohio State University as reported by President Rightmire (6) (p. 150) reveal a 5-10 per cent superiority in academic gains made by groups of students whose normal classroom work was supplemented by regular interviews as opposed to those who lacked this feature. Whether this slight but noticeable superiority is attributable to the *motivating* or the *informing* influence exerted by the interview—in which case it would merely be a disguised preceptorial system—is uncertain; the distinction between the two should be kept very clear and the relative effects of each submitted to measurement. If the interview is to be made a wholly beneficial adjunct to classroom instruction it must be considered in relation to the time factor, for unless a *substitute* hour of student interviewing results in greater personal progress than an equivalent teaching period its

efficiency is questionable—as is any other suggestion which might increase actual clock-hour loads. At the Pennsylvania State College some of the sections in educational psychology have recently been consolidated into larger groups (in accordance with Hudelson's (9) findings) and the lecture-recitation time thus saved has been converted into interview periods. The precise effects of this procedure are as yet unknown, but it is conceivable that where the Minnesota study merely showed an *equivalence* of large and small classes, this method may reveal that the large section plus individual interviews is definitely more advantageous on all counts. The interview occupies a most strategic position for procuring desirable attitudinal changes among individuals subjected to formal or informal educational processes. It is too early to report any definite outcomes but the illustration should suffice to call attention to the hidden possibilities which lurk within the conference method more widely applied.

REFERENCES

1. Bingham, W. V., & Moore, B. V., How to Interview, New York, 1931.
2. Charters, W. W., The collecting of unrecorded specifics, Journal of Educational Research, 1922, 5, 280-294.
3. Charters, W. W., Curriculum Construction, New York, 1923.
4. Clark, E. B., Value of student interviews, Journal of Personnel Research, 1926, 5, 204-207.
5. Davis, K. B., Factors in the Sex Life of Twenty-Two Hundred Women, New York, 1929.
6. Gray, W. S., ed., The training of college teachers, Chicago, 1930.
7. Hersey, R. B., Periodic emotional changes in male workers, Personnel Journal, 1929, 7, 459-464.
8. Houser, J. D., What the Employer Thinks, Cambridge, 1927.
9. Hudelson, E., Class size at the college level, Minneapolis, 1928.
10. O'Rourke, L. J., Measuring judgment and resourcefulness as interview technique, Personnel Journal, 1929, 7, 427-440.
11. Snedden, D. S., Disguised Intelligence Tests, Interview Form, Teachers College Contributions to Education, 1927, No. 291.

NEWS AND NOTES

Ever since the article entitled "Table of Standard Errors and Probable Errors of Percentages for Varying Numbers of Cases" by Harold A. Edgerton and Donald G. Paterson was published in this Journal in 1926 there has been a demand for reprints. We would now call the attention of the readers to the fact that we have had reprints made of this Table which can be obtained at the following rates: 30 cents for single copies; 20 cents each for orders of 5 to 24; 15 cents each for all orders of 25 or over. These may be secured from the following: James P. Porter, Editor, Ohio University, Athens, Ohio; Harold A. Edgerton, Department of Psychology, Ohio State University, Columbus, Ohio; or Donald G. Paterson, Department of Psychology, University of Minnesota, Minneapolis, Minn.

The attention of our readers is called to the article in the October, 1932, issue of the Journal of Applied Psychology by Dr. Harriet Babcock, the title being "The Short Army Performance Scale in Clinical Practice." It is the belief of the author and others that Miss Babcock's article will prove very useful to members of the staffs of psychological clinics. Inquiry concerning reprints may be addressed to the author, 25 Prospect Place, New York City, or to this Journal.

Donald A. Laird, director of the psychological laboratory of Colgate University, has been appointed to supervise the exhibit on sleep and fatigue for the Century of Progress Exposition, opening at Chicago in June. The exhibit, which is being supported by the Master Bedding Makers of America, will feature the actual use of apparatus used in recent studies of sleep and fatigue at several universities, giving visitors an opportunity to take some of the tests. It is expected to accumulate in this way a mass of data bearing on the diurnal course of fatigue.

Dr. Morris S. Viteles of the University of Pennsylvania, author of "Industrial Psychology," and Director of Personnel Research in the Philadelphia Electric Company, has been elected a member of the Board of Directors of the Personnel Research Federation.

With Dr. Goodwin Watson of the psychology department of Columbia University acting as director five stimulating psychology study-travel courses have been worked out for this season under the auspices of the American Peoples College in Europe. The courses are planned to meet the needs of psychology workers, students and teachers.

Four of the courses will include the Psycho-Technical Congress which convenes in Vienna from September 3 to 10. Other interesting features of the study tours will be the Psychological Institute of the University of Hamburg where a study will be made of Professor William Stern's theory of personality and the work of Professor Muchow with children and adolescents; the Psychoanalytic Institute, University of Berlin, Professor Moede's laboratory in the Technical High School, and the Institution of Juvenile Psychopathic Delinquents in Berlin, and a study of the work of Professor Volkeop in Leipsig and his theory of Gestalt Psychology. In Vienna members of the group will attend the Adler Psychological Clinics and the public school conducted according to the principle of individual psychology; Buhler's laboratory and demonstration of his psychological tests for children; the Juvenile Court; the Institution for Psychopathic Children; the Vienna Psychoanalytic Institute; and they will also study the work of Dr. Lazarsfeld in industrial psychology.

The Institute for Analytical Psychology will be visited in Zurich and an opportunity given for observation of the work of Dr. Jung. Paris offers a chance to study first-hand the clinic of Dr. Heuyer, observing the work with problem children in connection with the Juvenile Court; Professor Claude's hospital for mental diseases, and the laboratory of Dr. Simon of the Binet-Simon tests. Attendance at the DeCroly School in Brussels, one of the outstanding schools of the world, will give an opportunity to observe its theory and method of dealing with both normal and subnormal children. Further details of the travel-study courses may be obtained from Dr. Watson from the American headquarters of the American College in Europe at 55 West 42nd Street, New York City.

BOOK REVIEWS

GAULT, ROBERT H., *Criminology*, New York, Heath, 1932, pp. ix + 461. $3.48

The author in contrast with many others who have contributed to this field lays the greatest stress on attitude and personality. On the whole he seems to make his case. Another feature should be mentioned at the outset, namely, the extensive use of case studies to illustrate the various points which are made. Some of these cases are cited at considerable length and will be useful for teachers and lecturers in this field who wish well-chosen illustrative material without going to the original sources themselves. The book falls into two groups. The first on "the struggle against crime" is dedicated essentially to problems of prevention and immunization. The book is quite comprehensive in scope although some of the topics are touched on rather lightly but perhaps adequately for the purpose in hand.

Personality is defined as "behavior overt and implicit taken as a whole." This definition is certainly broad enough. The author criticises attempts to classify personality types. The chapter as a whole does not leave a very clear impression regarding personality.

The author is inclined to minimize the rôle of mental defect in delinquency except for repeaters. His conclusions are somewhat more conservative than those of a good many psychologists. Then follows a discussion of certain psychopathic symptoms and a few of the major psychoses with more emphasis on psychopathic personality, such as lying, malingering, and stealing. The author rightly feels that sex causation has been over-emphasized in such personalities. The only age factor of significance according to the author is the general uncertainty at adolescence.

He discusses the development of gangs and of the modern racket. The Jukes and Kallikak families are mentioned but the final conclusion is that the inheritance of mental defect, insanity and epilepsy according to Mendel's Law is rather doubtful.

The second part of the book is motivated by the thesis that the attitude of the offender is one of the main factors in the causation of crime. The psychology of the criminal during incarceration is handled rather well with due emphasis on the desirability of having him busy rather than idle and also teaching him some job which he may be able to follow subsequently. A wide variety of available occupations in prison is desirable. The prison personnel

can make or break the whole progress and efforts to train prison officers in special schools are most commendable. Stress is very properly laid on the selection of persons who ought to be paroled in the light of extensive study. Paying a fine in installments in lieu of imprisonment is suggested for certain types of cases.

Some 80 per cent of the offenders are not caught. The police force itself should be national rather than municipal in order to have the commanding authority far removed from local conditions and influences. There are numerous sources of error in observation and testimony. Psychological crime detection methods, such as the association reaction, breathing and blood pressure are covered briefly with commendable hesitance about hypnosis and scopolamine. Psychiatric and other experts should be summoned by the court and paid from court funds. He also commends the Massachusetts system of referring cases to the state hospital for observation.

The final chapter deals with prevention of the development of criminals, laying stress on the home, the school, and the juvenile court. The clinics in connection with courts are emphasized, likewise school and community clinics.

The reviewer is impressed by the number of references to the Journal of Criminal Law and Criminology and proposes to read that journal more regularly hereafter. The treatment is systematic and straightforward and will appeal to applied psychologists, to some persons in the legal field and also to sociologists and social workers. While it is not light reading in any sense it may even have some appeal to the more intelligent layman.

HAROLD E. BURTT,
Ohio State University.

SYMONDS, P. M.: *Diagnosing Personality and Conduct.* New York: The Century Co., 1931. Pp. xvi-602.

Treatises on psychological measurement have in the past dealt almost solely with such matters as intelligence, special aptitudes, and achievement in the various school subjects. But the school of today is not interested solely in the child's abilities or his success in subject matter. More important than anything that has yet been measured precisely is the adequacy of the child's personal adjustments. Within recent years, and particularly since the World War, research students in the field of measurement have been devoting a large portion of their time and effort to the study of character and personality. An unfortunate and perhaps inevitable feature of this endeavor is the fact that workers often have tended to arouse the expectation of practical applications of

their instruments that are at the present time quite impossible. Enthusiasm is perhaps essential if one is to work at all upon such an elusive problem as that of personality adjustment. Nevertheless it does seem that some workers are endowed with rather too much of the Pollyana attitude.

Diagnosing Personality and Conduct is a critical and a significant attempt to summarize and to evaluate the methods and the techniques that have been evolved by a host of individuals. This attempt has necessarily involved scrutiny and evaluation of data that have been assembled by the use of the available techniques. The first chapter of Professor Symonds's book discusses the significance of and the need for conduct diagnosis. Chapter II describes the general nature of the observational method and the limitations that inhere in diagnosis by the use of this method. Types of observation, the essentials of valid observation, the recording of observations, and the myriad pitfalls that are likely to be encountered by the unwary are also set forth. Adequacy of sampling, reliability and validity are likewise emphasized.

Chapter III is a rather lengthy treatment of rating methods. Not everyone will find this 80-page chapter particularly interesting but most readers will find it informing. Some of the things that receive especial emphasis in this chapter are:—the administration of ratings, general methods of making ratings, a comparison of various rating methods, an enumeration of the cardinal points that should be observed in rating, the statistical treatment of paired estimates and rankings, and the numerous factors that influence the validity and the reliability of ratings.

Chapters IV to VII inclusive present the most comprehensive treatment of the entire questionnaire problem, including the adjustment questionnaire, the attitude questionnaire, and the interest questionnaire, that the reviewer has seen in print. Chapter VIII is a cogent discussion of tests of knowledge, of conduct, and of judgment. Of particular interest in this chapter is the summary of those studies which have to do with the relationship between moral knowledge and moral conduct. Of these the author writes as follows:—

> "The testimony here is remarkably unanimous. . . . These correlations are very low. The conclusion is that there is a very low positive correlation between moral knowledge and desirable conduct." p. 289.

Chapter IX describes in detail tests of honesty, of suggestibility, of persistence, of cooperation or service, of inhibition, of caution, of speed of decision, of studiousness or effort in school, and the Downey Will-Temperament Tests. The author's candor in his ap-

praisal of these several devices is refreshing. Of Miss Downey's Will-Temperament he writes:—

"The chapter or section of a chapter in the history of psychology which will deal with the Downey 'Will-Temperament Tests' should be at the same time pathetic and amusing—pathetic because of the way in which gullible testers rushed in to use an instrument for which much had been promised, and amusing because of the politeness with which it was treated by scientific workers who had evidence of its worthlessness." p. 337.

"The general enthusiasm that originally greeted the Downey Will-Temperament Tests has subsided. Their practical value in guidance, personnel selection, and the like has not been demonstrated. Experimental work shows them to be tests of very specific abilities rather than tests to diagnose the more general personnel qualities the names of which they bear." p. 352.

It of course would be unfortunate to condemn a test before it has been given a thorough trial. On the other hand it seems equally unfortunate that scientists find it necessary to be so "polite." As the reviewer scanned the pages of Professor Symonds's book he could not help wondering how long it will be before the author's foregoing remarks will be equally applicable to certain other tests that are given generous, perhaps too generous, appraisal in this book.

Salient observations regarding the value and the limitations of the free association method as a method of diagnosis are presented in chapter X and the succeeding chapters have to do with the following topics:—physiological measures of the emotions, interviewing, psychoanalysis, the external signs of conduct, measures of the environment, and "The Case Study: A Comprehensive Study of the Individual."

Thorndike once stated that whatever exists exists in some amount and that "to measure it is simply to know its varying amounts." The foregoing statement has been quoted frequently by students of psychology and education and undoubtedly it has stimulated abundant effort on the part of all sorts of testers. Nevertheless, in spite of the salutary influence that this article of faith has had (and the foregoing statement should not be anything more than an article of faith) it is conceivable that, if adhered to too tenaciously or if interpreted too broadly, this pronouncement may become an impediment to future progress in the field of personality study. This of course is not the appropriate place for the reviewer to set forth his detailed views in this regard.* Suffice it to remark that some trait testers seem to have misinterpreted Thorndike's assertion. Some of the latter seem to

* Cf., in this connection H. C. Lehman and P. A. Witty, "'Faculty Psychology' and 'Personality Traits.'" *Psychological Bulletin.* (Forthcoming article.)

be laboring under the impression that "whatever exists exists in some amount," that "it can be measured," and that they "are able to measure it."

It is almost needless to remark that Professor Symonds does not make the last of the foregoing assumptions. He criticizes repeatedly those who assume:—(1) that the trait whose existence they posit must necessarily possess objective reality, and (2) that their tests of specific abilities actually measure the more general qualities that are implied by the names that they have given to their tests. On the whole the book under review is important both because of: (1) its comprehensiveness (almost 900 references are listed and 99 tables of data are presented), and (2) the sophisticated skepticism that is displayed by the author in his detailed and versatile analysis of the obstacles that must be surmounted if further progress in the field of personality measurement is to be attained. Most mature students of this subject are aware of the rudimentary character of most of the instruments that have thus far been evolved for the diagnosis of personality. Nevertheless, few of them who peruse this volume will fail to extend their acquaintance with the research that has been attempted in this field.

HARVEY C. LEHMAN,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Editor, *Journal of Applied Psychology*, Ohio University, Athens, Ohio.

Adjustment and Mastery. ROBERT S. WOODWORTH. Williams & Wilkins Company, Baltimore, Md., 1933. $1.00. 137 pp.

The Art of Educational Research. HAROLD H. ABELSON. World Book Company, Yonkers, N. Y., 1933. $2.00. 332 pp.

Discovering Ourselves. EDWARD A. STRECHER and KENNETH E. APPEL. Macmillan Company, New York City, 1932. $2.50. 306 pp.

Education for a New Era. A. GORDON MELVIN. John Day Company, New York City, 1933. 25 cents. 30 pp.

The First Two Years: A Study of Twenty-five Babies. Volume II. MARY M. SHIRLEY. University of Minnesota Press, Minneapolis, Minn., 1933. $3.00. 513 pp.

The Human Personality. LOUIS BERG. Prentice-Hall, Inc., New York City, 1933. 321 pp.

Instruction in English. DORA V. SMITH. Bulletin No. 17, 1932, Monograph No. 20. U. S. Department of the Interior, Washington, D. C. 10 cents. 89 pp.

Instruction in Science. U. S. Department of the Interior, Washington, D. C., Monograph No. 22, Bulletin No. 17. 63 pp.

Intelligence: Its Manifestations and Measurement. PAUL L. BOYNTON. D. Appleton & Company, New York City, 1933. $2.50. 466 pp.

Jugendpsychologie der Gegenwart. OTTO TUMBIRZ. Junker und Dunnhauft Verlag, Berlin-Steglitz. 97 pp.

The Measurement of Musical Development. HAROLD M. WILLIAMS, CLEMENT H. SIEVERS and MELVIN S. HATTWICK. University of Iowa Studies, Vol. 7, No. 1. University of Iowa, Iowa City, Iowa. 191 pp.

New York Spelling Tests, Series PW. Bureau of Reference, Research & Statistics, 1932. Publication No. 26. 16 pp. Board of Education, New York City.

New York Survey Tests in Arithmetic. Manual of Directions. Eugene A. Nifenecker, Director. Bureau of Reference, Research and Statistics. Board of Education. New York City. 16 pp.

Omaha Secret Societies. R. F. FORTUNE. Columbia University Contributions to Anthropology. Columbia University Press, New York City, 1932. $4.00. 193 pp.

The Performance of Student Nurses on Tests of Intelligence. ELSIE O. BREGMAN. Nursing Educational Bulletin, March, 1933. Department of Nurses' Education, Teachers College, Columbia University, New York City. 32 pp.

The Psychology of Childhood. NAOMI NORSWORTHY and MARY THEODORA WHITLEY. Macmillan Company, New York City, 1933. $1.80. 515 pp.

Psychological Principles of Education. FRANCIS F. POWERS and WILLIS L. UHL. The Century Company, New York City. 570 pp.

Quilente Texts. MANUEL J. ANDRADE. Columbia University Contributions to Anthropology, Vol. XII. Columbia University Press, New York City, 1931. $4.00. 211 pp.

Report of the President of Columbia University for 1932. Morningside Heights, N. Y. 50 pp.

Self-Instructional Manual in Handling Test Scores. RICHARD S. HARTER and C. H. SMELTZER. Henry Holt & Company, New York City, 1933. 58 pp.

The State and Higher Education. Carnegie Foundation for Advancement of Teaching and U. S. Office of Education, 1933. 522 Fifth Avenue, New York City. 282 pp.

Le Travail Humain. J. M. LAHY and H. LAUGEAR, Directors. Vol. 1, No. 1. Conservatoire National des Arts et Metiers. 292 Rue Saint-Marlin, Paris. 130 francs. 128 pp.

Directory of American Psychological Periodicals

(Continued)

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500-600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Philosophy—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monograph—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300-400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Indianapolis: The C. E. Pauley & Co., Inc. Subscription $5.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

(Continued on Third Cover)